

Ground State Energy of Bose Particle System

Ryuzo ABE

*Department of Physics, Tokyo Institute of Technology, Oh-okayama,
Meguro, Tokyo*

(Received July 22, 1958)

The ground state energy of the Bose particle system at low densities is investigated by using the method which is based on summing up the terms of the conventional perturbation series giving rise to the lowest order with respect to the density. The summation is shown to be greatly simplified if the scattering matrix is introduced, and a simple example is demonstrated for the repulsive square well potential by calculating a few terms of the series first directly and secondly with the aid of scattering matrix. The final result is shown to be valid both for weak and strong interactions and the calculation is carried out analytically for some types of potential function. The connection of the present method with the pseudopotential method and with the Brueckner theory is investigated and some problems associated with the attractive interactions are discussed.

§ 1. Introduction

In some problems of quantum mechanics or statistical mechanics a method is employed which is based on the summation of the important diagrams in the perturbation series¹⁻³⁾. In this paper we wish to investigate the ground state energy ε_0 per particle of the Bose system along this line.

The quantity ε_0 naturally depends on the density ρ of the system, on parameters characteristic of interatomic potential, etc. Now, let us imagine to plot ε_0 versus ρ curve and seek the behaviours of the curve near at $\rho=0$. At $\rho=0$ the system is essentially assembly of free particles and it follows $\varepsilon_0(\rho=0)=0$. Therefore it is sufficient to calculate $(\partial\varepsilon_0/\partial\rho)_{\rho=0}$ for our purpose. In other words, our problem is the idealized one to calculate the term linear in ρ when ε_0 is expanded in terms of ρ .

In § 2 the formal perturbation series in powers of the interatomic potential is discussed and the diagrams giving rise to the lowest order term of the ground state energy are identified to be those corresponding to the virtual processes in which the pair creations and annihilations take place successively. In § 3 the summation of the diagrams is carried out with the aid of the free-particle scattering matrix. A simple example of this procedure is demonstrated for the repulsive square well potential; first, a few terms of the series are calculated by direct integrations in § 2 and secondly by the use of scattering matrix in § 3. Both methods yield the same result for the terms of the series up to the order investigated in this paper. The lowest order term of the energy is calculated analytically for some types of the interatomic potential, and the problem associated with the presence of the bound states is discussed. In § 4 the perturbation series in terms of the

interatomic potential is transformed into the one in terms of the scattering matrix and it is shown that the diagrams summed up in § 2 do not appear in the new series and that an application to the hard-sphere potential leads to an expression obtained by pseudo-potential method. The alteration of the energy denominator proposed by Brueckner and Sawada⁴⁾ is interpreted in an elementary way from the view-point of the present paper. It is suggested that an appropriate variational procedure is necessary to deal with the Bose system with attractive but saturating interactions.

§ 2. Perturbation series for the Bose system

Let us discuss in this section the behaviour of the formal perturbation series in powers of the interatomic potential for the Bose system. We assume that the unperturbed Hamiltonian is

$$\mathcal{H}_0 = \hbar^2/2m \cdot \sum_{\alpha} k_{\alpha}^2 a_{\alpha}^* a_{\alpha} \quad (1)$$

and the perturbing Hamiltonian is

$$\mathcal{H}' = \hbar^2/2m \cdot 1/2V \cdot \sum_{\alpha\beta\lambda\gamma} v(\alpha\beta, \lambda\gamma) a_{\alpha}^* a_{\beta}^* a_{\gamma} a_{\lambda}. \quad (2)$$

The operators a^* and a are the usual creation and annihilation operators for the Bose particles. The matrix elements $v(\alpha\beta, \lambda\gamma)$ are taken in the momentum representation, i.e.,

$$\begin{aligned} v(\alpha\beta, \lambda\gamma) &= \frac{1}{V} \int \exp(-ik_{\alpha} \cdot x_1 - ik_{\beta} \cdot x_2) v(1,2) \exp(ik_{\lambda} \cdot x_1 + ik_{\gamma} \cdot x_2) dx_1 dx_2 \\ &= \delta(k_{\alpha} + k_{\beta} - k_{\gamma} - k_{\lambda}) \nu(k_{\alpha\beta}, k_{\lambda\gamma}) \end{aligned} \quad (3)$$

where V is the normalization volume, $\hbar^2/2m \cdot v(1,2)$ the two-body potential, $k_{\alpha\beta} = (k_{\alpha} - k_{\beta})/2$, $k_{\lambda\gamma} = (k_{\lambda} - k_{\gamma})/2$, and $\nu(q, q')$ is defined by

$$\nu(q, q') = \int \exp(-iq \cdot x) v(x) \exp(iq' \cdot x) dx. \quad (4)$$

In the following we shall consider the non-singular two-body potentials so that the perturbation method can be applied. To deal with the case of a singular potential such as a hard-sphere one, it is still convenient to work directly with the matrix elements $v(\alpha\beta, \lambda\gamma)$ by adopting some appropriate cutoff procedure. We shall prove later on that the final result is independent of the choice of such procedures.

Now, each term of the perturbation series can be written as an integral over various vectors q_i , which are virtual momentum transfers. The successive terms of the series are:

$$\begin{aligned} E_1 &= \langle 0 | \mathcal{H}' | 0 \rangle = \hbar^2/2m \cdot N(N-1)/2V \cdot v(00, 00) \\ &= \hbar^2/2m \cdot N(N-1)/2V \cdot \nu(0, 0) \end{aligned} \quad (5)$$

where N is the total number of particles. The second-order energy is

$$\begin{aligned}
 E_2 &= \langle 0 | \mathcal{H}' \frac{1}{-\mathcal{H}_0} \mathcal{H}' | 0 \rangle = \frac{\hbar^2}{2m} \frac{N(N-1)}{(2V)^2} \sum_{q \neq 0} \frac{v(00, -\mathbf{q}\mathbf{q}) + v(00, \mathbf{q}-\mathbf{q})}{(-2q^2)} v(\mathbf{q}-\mathbf{q}, 00) \\
 &= \frac{\hbar^2}{2m} \frac{N(N-1)}{(2V)^2} \sum_{q \neq 0} \frac{\nu(0, \mathbf{q}) \nu(\mathbf{q}, 0)}{(-q^2)} \quad (6)
 \end{aligned}$$

where use is made of eq. (3). In order to proceed to the higher order terms, it is convenient to introduce a graphical representation of the virtual processes as was used by Brueckner and Sawada⁴. In Fig. 1, we give the diagram which represents a process giving rise to E_2 .

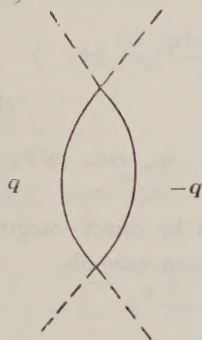


Fig. 1 Diagram contributing to E_2 . The excited particles are indicated by solid lines and the unexcited particles by dashed lines.

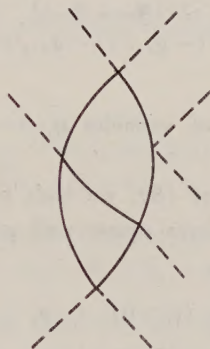


Fig. 2

Then, the diagrams corresponding to the processes that lead to the term linear in ρ are easily identified. Consider, for example, the diagram shown in Fig. 2. The process corresponding to this is characterized by the annihilation or creation of the unexcited

particles (particles with 0-momentum) in the intermediate states and therefore the energy term contains a factor $\langle 0 | a_0^* a_0^* a_0^* a_0^* a_0 a_0 a_0 | 0 \rangle \sim N^4$ in this case. Now the energy should be proportional to N and this suggests that the term under consideration has the form $N\rho^3$ and is of the higher orders with respect to ρ (the factor V^{-3} comes from \mathcal{H}' and the sum over intermediate momentum states). We see from this argument that the lowest order term in ρ is such that the annihilation or creation of the unexcited particles must not occur except for the initial or final state. Then the diagrams having this character are easily written down. Since the annihilation of the unexcited particles does not occur in the intermediate states, in each stage of the virtual processes the pair with momenta \mathbf{q} and $-\mathbf{q}$ should be destroyed and excited to the states $(\mathbf{q}', -\mathbf{q}')$, etc., and finally returns to the ground state. Thus the diagrams interesting to us will be of the type shown in Fig. 3. Then the energy term in the n -th order corresponding to this process is easily shown to be

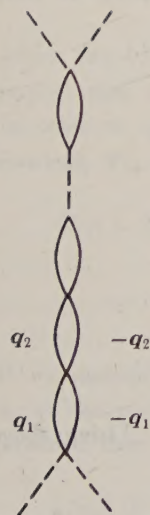


Fig. 3

$$\begin{aligned}
 &\frac{\hbar^2 N(N-1)}{2m(2V)^n} \sum_{q_1, \dots, q_{n-1}} \frac{\{v(00, \mathbf{q}_{n-1} - \mathbf{q}_{n-1}) + v(00, -\mathbf{q}_{n-1} \mathbf{q}_{n-1})\} \cdots \{v(\mathbf{q}_2 - \mathbf{q}_2, \mathbf{q}_1 - \mathbf{q}_1) + \\
 &\quad v(\mathbf{q}_2 - \mathbf{q}_2, -\mathbf{q}_1 \mathbf{q}_1)\} v(\mathbf{q}_1 - \mathbf{q}_1, 00)}{(-2q_{n-1}^2)(-2q_{n-2}^2) \cdots (-2q_1^2)} \quad (7)
 \end{aligned}$$

where \sum' means that $q_1 \neq 0, q_2 \neq 0, \dots, q_{n-1} \neq 0$. Therefore, if we define E_0 to be the lowest order term of the energy, we have

$$E_0 = \frac{\hbar^2 N(N-1)}{4mV} \left\{ \nu(0, 0) - \frac{1}{(2V)} \sum'_{q_1} \frac{\nu(0, q_1) \nu(q_1, 0)}{q_1^2} \right. \\ + \frac{1}{(2V)^2} \sum'_{q_1 q_2} \frac{\nu(0, q_2) \nu(q_2, q_1) \nu(q_1, 0)}{q_2^2 q_1^2} \\ \left. + \frac{1}{(2V)^{n-1}} \sum'_{q_1, \dots, q_{n-1}} \frac{\nu(0, q_{n-1}) \nu(q_{n-1}, q_{n-2}) \dots \nu(q_2, q_1) \nu(q_1, 0)}{(-q_{n-1}^2) (-q_{n-2}^2) \dots (-q_1^2)} + \dots \right\}. \quad (8)$$

Here we have carried out the change of variables $q_1 \rightarrow -q_1, q_2 \rightarrow -q_2$, etc., in eq. (7) and used eq. (3).

Before discussing the summation of (8), we wish to calculate by direct integrations a few terms of the series for the repulsive square well potential as an example.

For the potential

$$v(x) = v_0 \quad (v_0 > 0), \quad r < a, \quad (r = |x|), \\ = 0, \quad r > a, \quad (9)$$

$\nu(q, q')$ is given by

$$\nu(q, q') = 4\pi v_0 a^2 j_1(ka) / k \quad (10)$$

where $j_1(x)$ is the spherical Bessel function of order 1 and $k = |q - q'|$. Then we have

$$\nu(0, 0) = 4\pi v_0 a^3 / 3, \\ \frac{1}{(2V)} \sum'_{q_1} \frac{\nu(0, q_1) \nu(q_1, 0)}{q_1^2} = 4\pi v_0 a^4 \int_0^\infty \frac{j_1^2(qa)}{q^2} dq = \frac{4\pi v_0 a^5}{15}$$

and

$$\frac{1}{(2V)^2} \sum'_{q_1 q_2} \frac{\nu(0, q_1) \nu(q_2, q_1) \nu(q_2, 0)}{q_1^2 q_2^2} \\ = \frac{2\pi v_0 a^4}{\pi} \int_0^\infty \frac{j_1(q_1 a)}{q_1^2} dq_1 \int_{-\infty}^\infty \frac{j_1(q_2 a) \sin a(q_1 - q_2)}{q_2^2 (q_1 - q_2)} dq_2 \\ = 4\pi v_0 a^7 \cdot 17 / 1260.$$

Here we have used the Weber-Schafheitlin integral for the Bessel function⁵⁾. Using above equations, we have from eq. (8)

$$\frac{E_0}{N} = \frac{2\pi \hbar^2 \rho}{m} \left\{ \frac{v_0 a^3}{6} - \frac{v_0^2 a^5}{30} + \frac{17 v_0^3 a^7}{2520} - \dots \right\} \quad (11)$$

up to the order of v_0^3 . The further calculation is cumbersome and it is difficult to find the general term of the series. A convenient method to calculate the series will be given in the next section.

§ 3. Summation of the series

In order to sum up the series (8), it is convenient to replace the problem of summation by an equivalent and a simpler one. For this purpose we wish to proceed to the more general case where the unperturbed state is represented by the occupation numbers n_k 's, not necessarily corresponding to the ground state. Then it is easily verified that the energy term arising from the diagrams discussed in § 2 is given by

$$E_0(n_k) = \frac{\hbar^2}{2m} \frac{1}{2V} \sum_{\alpha\beta\lambda\tau} u(\alpha\beta, \lambda\tau) \langle n_k | a_\alpha^* a_\beta^* a_\tau a_\lambda | n_k \rangle \quad (12)$$

where

$$\begin{aligned} u(\alpha\beta, \lambda\tau) = & v(\alpha\beta, \lambda\tau) + \frac{1}{V} \sum'_{\alpha_1\beta_1} \frac{v(\alpha\beta, \alpha_1\beta_1) v(\alpha_1\beta_1, \lambda\tau)}{k_\lambda^2 + k_\tau^2 - k_{\alpha_1}^2 - k_{\beta_1}^2} \\ & + \frac{1}{V^2} \sum'_{\substack{\alpha_2\beta_2 \\ \alpha_1\beta_1}} \frac{v(\alpha\beta, \alpha_2\beta_2) v(\alpha_2\beta_2, \alpha_1\beta_1) v(\alpha_1\beta_1, \lambda\tau)}{(k_\lambda^2 + k_\tau^2 - k_{\alpha_2}^2 - k_{\beta_2}^2) (k_\lambda^2 + k_\tau^2 - k_{\alpha_1}^2 - k_{\beta_1}^2)} + \dots \end{aligned} \quad (13)$$

In eq. (13) \sum' means that the values of k_{α_1} , k_{β_1} , etc., which make the energy denominator vanish, are excluded from the summation. One may verify that eq. (13) reduces to eq. (8) in the case of ground state, i.e. $n_0 = N$ and $n_k = 0$ ($k \neq 0$).

Now, we note that the matrix elements $u(\alpha\beta, \lambda\tau)$ satisfy the equation

$$v(\alpha\beta, \lambda\tau) + \frac{1}{V} \sum'_{\sigma\tau} v(\alpha\beta, \sigma\tau) \frac{u(\sigma\tau, \lambda\tau)}{k_\lambda^2 + k_\tau^2 - k_\sigma^2 - k_\tau^2} = u(\alpha\beta, \lambda\tau), \quad (14)$$

which can be proved from the defining equation for $u(\alpha\beta, \lambda\tau)$, eq. (13). This equation implies that $u(\alpha\beta, \lambda\tau)$ is just the matrix element of the free-particle scattering operator⁽⁶⁾. In order to transform this equation in the coordinate representation, we consider the wave function $\mathcal{F}_{\lambda\tau}(1, 2)$ defined by

$$\begin{aligned} \mathcal{F}_{\lambda\tau}(1, 2) = & \exp(i\mathbf{k}_\lambda \cdot \mathbf{x}_1 + i\mathbf{k}_\tau \cdot \mathbf{x}_2) + \frac{1}{V^2} \sum'_{\sigma\tau} \frac{\exp(i\mathbf{k}_\sigma \cdot \mathbf{x}_1 + i\mathbf{k}_\tau \cdot \mathbf{x}_2)}{k_\lambda^2 + k_\tau^2 - k_\sigma^2 - k_\tau^2} \\ & \times \int \exp(-i\mathbf{k}_\sigma \cdot \mathbf{x}'_1 - i\mathbf{k}_\tau \cdot \mathbf{x}'_2) v(1', 2') \mathcal{F}_{\lambda\tau}(1', 2') d\mathbf{x}'_1 d\mathbf{x}'_2. \end{aligned} \quad (15)$$

If we multiply this equation by $\exp(-i\mathbf{k}_\alpha \cdot \mathbf{x}_1 - i\mathbf{k}_\beta \cdot \mathbf{x}_2) v(1, 2)$ and integrate by \mathbf{x}_1 and \mathbf{x}_2 in the volume V , we see that the resulting equation is just identical with eq. (14) provided that $u(\alpha\beta, \lambda\tau)$ is given by

$$u(\alpha\beta, \lambda\tau) = \frac{1}{V} \int \exp(-i\mathbf{k}_\alpha \cdot \mathbf{x}_1 - i\mathbf{k}_\beta \cdot \mathbf{x}_2) v(1, 2) \mathcal{F}_{\lambda\tau}(1, 2) d\mathbf{x}_1 d\mathbf{x}_2. \quad (16)$$

Or, if we put

$$\mathcal{F}_{\lambda\tau}(1, 2) = \exp[i(\mathbf{k}_\lambda + \mathbf{k}_\tau) \cdot \mathbf{X}] \varphi_{\lambda\tau}(\mathbf{x})$$

where \mathbf{X} is the center-of-mass coordinate $[=(\mathbf{x}_1 + \mathbf{x}_2)/2]$ and \mathbf{x} the relative coordinate

($=\mathbf{x}_1-\mathbf{x}_2$), then we have

$$u(\alpha\beta, \lambda_T) = \delta(\mathbf{k}_\alpha + \mathbf{k}_\beta - \mathbf{k}_\lambda - \mathbf{k}_T) \int \exp(-i\mathbf{k}_{\alpha\beta} \cdot \mathbf{x}) v(\mathbf{x}) \varphi_{\lambda_T}(\mathbf{x}) d\mathbf{x} \quad (17)$$

and

$$\begin{aligned} \varphi_{\lambda_T}(\mathbf{x}) = & \exp(i\mathbf{k}_{\lambda_T} \cdot \mathbf{x}) \\ & + \frac{1}{2V} \sum' \frac{\exp(i\mathbf{k}_{\sigma_T} \cdot \mathbf{x})}{k_{\lambda_T}^2 - k_{\sigma_T}^2} \int \exp(-i\mathbf{k}_{\sigma_T} \cdot \mathbf{x}') v(\mathbf{x}') \varphi_{\lambda_T}(\mathbf{x}') d\mathbf{x}'. \end{aligned} \quad (18)$$

If we operate ∇^2 on eq. (18), we have

$$\left[-\nabla^2 + \frac{v(\mathbf{x})}{2} \right] \varphi(\mathbf{x}) = k_{\lambda_T}^2 \varphi(\mathbf{x}) + \frac{1}{2V} \sum_{|\mathbf{q}|=k_{\lambda_T}} \int \exp[i\mathbf{q} \cdot (\mathbf{x} - \mathbf{x}')] v(\mathbf{x}') \varphi(\mathbf{x}') d\mathbf{x}'. \quad (19)$$

This is the Schrödinger equation for the scattering problem except for the second term on the right-hand side. The appearance of this term is due to the finite volume of the system and consistent with the periodic boundary condition imposed on $\varphi(\mathbf{x})$. This will be seen in the following way; if we multiply eq. (19) by $\exp(-i\mathbf{q}_0 \cdot \mathbf{x})$ with \mathbf{q}_0 satisfying the periodic condition, we have

$$(q_0^2 - k_{\lambda_T}^2) \int \exp(-i\mathbf{q}_0 \cdot \mathbf{x}) \varphi(\mathbf{x}) d\mathbf{x} = 1/2 \cdot [\delta(\mathbf{k}_{\lambda_T}, \mathbf{q}_0) - 1] \int \exp(-i\mathbf{q}_0 \cdot \mathbf{x}) v(\mathbf{x}) \varphi(\mathbf{x}) d\mathbf{x}.$$

From this equation it follows that the appearance of the second term is consistent with the boundary condition, since, if this term is missing in eq. (19), then it follows $\int \exp(-i\mathbf{q}_0 \cdot \mathbf{x}) v(\mathbf{x}) \varphi(\mathbf{x}) d\mathbf{x} = 0$, for $\mathbf{q}_0 = \mathbf{k}_{\lambda_T}$, which is not always valid for an arbitrary value of k_{λ_T} satisfying the periodic condition. In the limit $V \rightarrow \infty$, however, this term may be neglected and eq. (19) reduces to the usual scattering equation.

Returning now to the summation of (8), we see that E_0 is given by

$$E_0/N = \hbar^2 \rho / 4m \cdot u(00,00). \quad (20)$$

In calculating $u(00,00)$ it is sufficient to consider only the spherically symmetric solution of eq. (19). Also the solution should approach 1 for $r \equiv |\mathbf{x}| \rightarrow \infty$, for $\varphi(\mathbf{x}) \rightarrow \exp(i\mathbf{k}_{\lambda_T} \cdot \mathbf{x})$ at large r and $\mathbf{k}_{\lambda_T} = 0$ in this case. Furthermore, the second term on the right-hand side of eq. (19) may be neglected as we have mentioned above.

From these considerations it follows that

$$u(00,00) = 4\pi \int_0^\infty v(r) R(r) r^2 dr \quad (21)$$

where $R(r)$ satisfies the equations

$$[1/r^2 \cdot d/dr(r^2 d/dr) - v(r)/2] R(r) = 0 \quad (22a)$$

and

$$R(r) \rightarrow 1, \quad (r \rightarrow \infty). \quad (22b)$$

Furthermore, $R(r)$ should be regular at $r=0$.

From eqs. (20), (21) and (22a) we have

$$E_0/N = 2\pi\hbar^2\rho/m \cdot (r^2 dR/dr)_{r=\infty}. \quad (23)$$

We see that eqs. (22a), (22b) and (23) are applicable to the strong interactions and identical with those obtained in the previous paper⁷⁾. For the repulsive square well potential considered in § 2, the solution of eq. (22a) is

$$R=1+A/r, \quad (r>a) \quad \text{and} \quad R=B \operatorname{sh}\sqrt{v_0/2} \cdot r/r, \quad (r<a).$$

From the condition that $R(r)$ and $R'(r)$ should be continuous at $r=a$, we have

$$A=\sqrt{2/v_0} \operatorname{th} a\sqrt{v_0/2} - a.$$

Therefore from eq. (23), we have

$$E_0/N = 2\pi\hbar^2\rho/m \cdot \{a - \sqrt{2/v_0} \operatorname{th} a\sqrt{v_0/2}\}. \quad (24)$$

If we expand $\operatorname{th} x$ as

$$\operatorname{th} x = x - 1/3 \cdot x^3 + 2/15 \cdot x^5 - 17/315 \cdot x^7 + \dots$$

the result is easily shown to be identical with eq. (11) obtained by direct integrations.

It should be noted here that eq. (24) remains finite as $v_0 \rightarrow \infty$. On the other hand, if $\operatorname{th} x$ is expanded in powers of x , each term of the series is divergent in this limit. This feature of the case of strong interaction implies that an appropriate procedure of summing up all the divergent terms of the conventional perturbation series is necessary to obtain a finite result. For the potential under consideration, the limiting case is just the hard-sphere one, and eq. (24) is naturally reduced to the formula obtained by Lenz⁸⁾. We wish further to remark that the discussions developed in this paper are also applicable to the weak interactions for which $\nu(\mathbf{q}, \mathbf{q}')$ is finite. In this case, each term of the perturbation series may be finite. However, the result is meaningless if the series is terminated in the finite order terms as far as the density dependence of the energy is concerned. Therefore it follows that the summation of diagrams is necessary even for weak interactions if one wishes to investigate a correct functional relation of the energy with the density.

In conclusion we wish to calculate E_0/N for some types of potential function.

(a) Attractive square well potential with the hard-core repulsion

In this case $v(\mathbf{x})$ is given by

$$\begin{cases} v(\mathbf{x}) = \infty, & r < a, \\ v(\mathbf{x}) = -v_0, & (v_0 > 0), \quad a < r < a+d, \\ v(\mathbf{x}) = 0, & a+d < r. \end{cases}$$

It is easy to see that

$$\frac{E_0}{N} = \frac{2\pi\hbar^2\rho}{m} [a+d - \sqrt{2/v_0} \tan \sqrt{v_0/2} \cdot d]. \quad (25)$$

(b) Sutherland potential

Two-body potential $V(r)$ is defined by

$$V(r) = \infty, \quad r < a,$$

$$V(r) = -V_0 (a/r)^n, \quad (V_0 > 0), \quad r > a.$$

For this potential the solution of (22a) is given by the Bessel function for $r > a$, and the boundary conditions at $r = \infty$, eq. (22b), and at $r = a$, $R(a) = 0$, uniquely determine the solution. Then we have

$$\frac{E_0}{N} = \frac{2\pi\hbar^2\rho}{m} \frac{\Gamma\{(n-3)/(n-2)\}}{\Gamma\{(n-1)/(n-2)\}} \left(\frac{k}{n-2}\right)^{2/(n-2)} \frac{J_{1/(2-n)}\{2ka^{(2-n)/2}/(n-2)\}}{J_{1/(n-2)}\{2ka^{(2-n)/2}/(n-2)\}} \quad (26)$$

with $k^2 = mV_0 a^n / \hbar^2$.

In particular, for the van der Waals attraction, $n=6$, eq. (26) reduces to

$$\frac{E_0}{N} = \frac{2\pi\hbar^2\rho}{m} \frac{\Gamma(3/4)}{\Gamma(5/4)} \frac{1}{2} \left(\frac{mV_0}{\hbar^2}\right)^{1/4} a^{3/2} \frac{J_{-1/4}\{(mV_0)^{1/2} a / 2\hbar\}}{J_{1/4}\{(mV_0)^{1/2} a / 2\hbar\}}. \quad (26a)$$

In the limit $V_0 \rightarrow 0$, eq. (26a) becomes Lenz's formula⁹⁾ for the hard-sphere potential.

(c) Lennard-Jones (6-10) potential

This potential was discussed in the previous paper⁷⁾, so that we only give the result:

$$\frac{E_0}{N} = 4\pi \frac{2^{3/4}\Gamma(3/4)\Gamma(5/8-\eta)}{\Gamma(1/4)\Gamma(3/8-\eta)} m^{-7/8} V_0^{1/8} \sigma^{5/4} \hbar^{7/4} \rho. \quad (27)$$

It should be noted here that the E_0/N becomes $-\infty$ if the bound state with zero energy appears in the two-body system. For example, if $\sqrt{(v_0/2)d}$ is equal to $\pi/2$ in eq. (25), $E_0/N = -\infty$. The above condition for v_0 and d just corresponds to the appearance of the bound state with zero energy. In this case E_0/N versus ρ curve coincides with the negative E_0/N axis near $\rho=0$. The similar situation exists for the potential functions (b) and (c). In the case of (b), (we take $n=6$ for simplicity), as V_0 increases from 0, E_0/N decreases from the positive value until it becomes $-\infty$ when $(mV_0)^{1/2} a / 2\hbar = \hat{\epsilon}$, where $\hat{\epsilon}$ is the least positive root of $J_{1/4}(\hat{\epsilon}) = 0$. Similarly, for the potential (c), E_0/N becomes $-\infty$ at $\eta = 5/8$. It seems, therefore, that the behaviour of the energy versus density curve is markedly dependent on the presence of the bound state in the two-body problem.

§ 4. Perturbation method in terms of scattering matrix

In the previous sections we have discussed the summation of the diagrams giving rise to the lowest order term of the energy. If one compares eq. (5) with eq. (20), one may easily see that E_0 is just the first order energy with $v(\alpha\beta, \lambda\gamma)$ replaced by $u(\alpha\beta, \lambda\gamma)$. This suggests that the potential function is effectively replaced by the scattering matrix. In this section we wish to discuss the perturbational treatment of the

scattering matrix.

Let us first consider the meaning of eq. (14). Originally this equation was introduced for the purpose of summing up the series (8) and u -matrix is explicitly expanded in terms of $\nu(q', q)$ as in eq. (13). However, we may regard eq. (14) conversely as the defining equation for $\nu(q', q)$ in terms of u -matrix. Then we have, writing $u(\alpha\beta, \lambda\gamma) = u(q', q)$ with $k_{\alpha\beta} = q'$ and $k_{\lambda\gamma} = q$ [see eq. (17)],

$$\begin{aligned} \nu(q', q) = & u(q', q) + \frac{1}{2V} \sum_{q''} \frac{u(q', q'') u(q'', q)}{q''^2 - q^2} \\ & + \frac{1}{(2V)^2} \sum_{q'' q'''} \frac{u(q', q'') u(q'', q''') u(q''', q)}{(q''^2 - q^2)(q'''^2 - q''^2)} + \dots \end{aligned} \quad (28)$$

If one substitutes this equation in the perturbation series in powers of ν -matrix, the series will be transformed in the one in powers of scattering matrix. For example, if one retains only the first term in eq. (28), the first order energy will be given by eq. (20). As is well known, the scattering matrix does exist even for the strong interaction and therefore it may be expected that the series in terms of u -matrix is valid for such a case. Furthermore, the following feature of the series may be expected: since the diagrams, in which the creation or annihilation of the unexcited particles does not occur except for the initial or final state, are included already in the u -matrix, such diagrams will not appear in the u -matrix series. In other words, only the diagrams corresponding to the process of the creation or annihilation of the unexcited particles in the intermediate states will appear in the new series. That this is the case will be verified up to the third order in the following.

The first and the second order energy were given by eqs. (5) and (6) and the third order one is easily shown to be

$$\begin{aligned} E_3 = & \frac{\hbar^2}{2m} \frac{N(N-1)}{(2V)^3} \left[\sum_{q'} \frac{\nu(0, q') \nu(q', q) \nu(q, 0)}{q^2 q'^2} + \sum_q \frac{\nu(0, q) \nu(q, 0)}{q^4} \right. \\ & \times \{ 2(N-2) [\nu(q/2, q/2) + \nu(-q/2, q/2)] - (2N-3) \nu(0, 0) \} \Big]. \end{aligned} \quad (29)$$

If we substitute eq. (28) in this expression and in eqs. (5) and (6) and express the perturbation series in terms u -matrix, we have

$$\begin{aligned} E = & \frac{\hbar^2}{2m} \frac{N(N-1)}{2V} \left[u(0, 0) + \frac{1}{(2V)^2} \sum_q \frac{u(0, q) u(q, 0)}{q^4} \right. \\ & \times \{ 2(N-2) [u(q/2, q/2) + u(-q/2, q/2)] - (2N-3) u(0, 0) \} \Big] \end{aligned} \quad (30)$$

up to the third order. It should be noted here that the diagrams corresponding to Fig. 3 no longer appear in eq. (30), as we have expected. The similar feature of the new series may be true for any higher order terms.

If we consider the hard-sphere system and set $u(0, 0)$ to be $8\pi a$, neglecting the dependence of u -matrix on q , we have from eq. (30)

$$\frac{E}{N} = \frac{2\pi a \hbar^2 \rho}{m} \left[1 + \frac{a^2}{\pi^2 L^2} (2N-5) \sum' \frac{1}{(l^2 + m^2 + n^2)^2} \right]$$

($V=L^3$), which is consistent with the result obtained by Huang and Yang⁹⁾ with the aid of pseudopotential method. Therefore we are able to conclude that the pseudopotential method is equivalent to the perturbation theoretical treatment of the free-particle scattering matrix. However, as was pointed out by Brueckner and Sawada⁴⁾, the perturbation method is not appropriate in this case, for the energy term corresponding to the diagram shown in Fig. 4 is divergent. To avoid this difficulty they have summed up the diagrams of this type up to infinite order. This process seems to be quite reasonable if

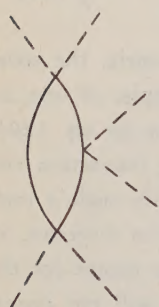


Fig. 4 Divergent diagram.

the u -matrix formalism is employed, since the diagrams appearing in the new series must at least once create or annihilate the unexcited particles in the intermediate states. Indeed, as we have shown previously¹⁰⁾, this formalism leads to the exact formula obtained by Lee, Huang and Yang¹¹⁾ which is appropriate for the low density limits. Therefore one may conclude that the free-particle scattering matrix formalism is exact as long as the density of the system is sufficiently low.

At higher densities, however, the free-particle scattering theory will become poor, since, in this case, the many-body effects of the unexcited particles on the virtual pair will not be neglected. In order to deal with these effects, Brueckner and Sawada⁴⁾ have altered the energy denominator so that the Green function describes the propagation in the presence of the unexcited particles. This propagator alteration may be interpreted in the following way. Also this procedure will provide a clue to solving the difficulties associated with the attractive interactions.

As we have done in § 2, we start from the weak interaction, i.e., we consider the Hamiltonian

$$\mathcal{H} = \frac{\hbar^2}{2m} \sum_{\alpha} k_{\alpha}^2 a_{\alpha}^* a_{\alpha} + \frac{\hbar^2}{2m} \frac{1}{2V} \sum_{\alpha\beta\lambda\gamma} v(\alpha\beta, \lambda\gamma) a_{\alpha}^* a_{\beta}^* a_{\gamma} a_{\lambda}, \quad (31)$$

where $v(\alpha\beta, \lambda\gamma)$ is assumed to have a definite value. In § 2, the first term (kinetic energy) was considered as the unperturbed Hamiltonian and the second term as the perturbing one, and the multiple interaction of a pair of particles excited from the ground state was treated. Then we were led to the concept of free-particle scattering matrix. Hence, it will be necessary to choose another way of separation of the Hamiltonian into the unperturbed and the perturbing one, if one wishes to take into account the effects of the unexcited particles on the propagator. The simplest way of this separation will be such that the parts of the interaction Hamiltonian which are diagonal in the number representation are included in the unperturbed system. That is to say, the unperturbed Hamiltonian is considered to be

$$\mathcal{H}_0 = \frac{\hbar^2}{2m} \left[\sum_{\alpha} k_{\alpha}^2 a_{\alpha}^* a_{\alpha} + \frac{1}{2V} \sum_{\alpha \neq \beta} \{v(\alpha\beta, \alpha\beta) + v(\alpha\beta, \beta\alpha)\} a_{\alpha}^* a_{\beta}^* a_{\alpha} a_{\beta} \right]$$

$$+\frac{1}{2V}\sum_{\alpha}v(\alpha\alpha, \alpha\alpha)a_{\alpha}^*a_{\alpha}^*a_{\alpha}a_{\alpha}\Big] \quad (32)$$

and the remaining part is assumed to be the perturbing one.

Now consider the unperturbed wave vector $|n_0, n_1, \dots, n_k, \dots\rangle$. The unperturbed energy corresponding to this is

$$E_0(n_k) = \frac{\hbar^2}{2m} \left[\sum_{\alpha} k_{\alpha}^2 n_{\alpha} + \frac{1}{2V} \sum_{\alpha \neq \beta} \{v(\alpha\beta, \alpha\beta) + v(\alpha\beta, \beta\alpha)\} n_{\alpha} n_{\beta} \right. \\ \left. + \frac{1}{2V} \sum v(\alpha\alpha, \alpha\alpha) n_{\alpha} (n_{\alpha} - 1) \right] \quad (33)$$

Therefore if one repeats the procedure discussed in § 2 and § 3, the energy denominator in eq. (13) will be altered in the following way:

$$\hbar^2 \cdot 2m \cdot (k_{\gamma}^2 + k_{\gamma}^2 - k_{\alpha}^2 - k_{\beta}^2) \rightarrow E_0(n_0, n_1, \dots) - E_0(n_0, \dots, n_{\alpha} + 1, n_{\beta} + 1, n_{\lambda} - 1, n_{\gamma} - 1) \\ = \frac{\hbar^2}{2m} \{ \varepsilon(\lambda) + \varepsilon(\gamma) - \varepsilon(\alpha) - \varepsilon(\beta) \} \quad (34)$$

where

$$\varepsilon(\lambda) = k_{\lambda}^2 + \frac{1}{V} \sum_{\lambda'} \{v(\lambda\lambda', \lambda\lambda') + v(\lambda\lambda', \lambda'\lambda)\} n_{\lambda'} - \frac{1}{V} v(\lambda\lambda, \lambda\lambda) n_{\lambda}. \quad (35)$$

As we have remarked in the previous sections, the matrix elements $v(\alpha\beta, \lambda\gamma)$ can be replaced by $u(\alpha\beta, \lambda\gamma)$ in the case of strong interactions. Therefore, if we replace v 's in eq. (35) by u 's, using eq. (14) we have

$$v(\alpha\beta, \lambda\gamma) + \frac{1}{V} \sum_{\sigma\tau} v(\alpha\beta, \sigma\tau) \frac{u(\sigma\tau, \lambda\gamma)}{\varepsilon(\lambda) + \varepsilon(\gamma) - \varepsilon(\sigma) - \varepsilon(\tau)} = u(\alpha\beta, \lambda\gamma) \quad (36)$$

as an equation to determine the u -matrix. Here $\varepsilon(\lambda)$ corresponds to the energy of a particle interacting with the unexcited particles and is given by

$$\varepsilon(\lambda) = k_{\lambda}^2 + \frac{1}{V} \sum_{\lambda'} \{u(\lambda\lambda', \lambda\lambda') + u(\lambda\lambda', \lambda'\lambda)\} n_{\lambda'} - \frac{1}{V} u(\lambda\lambda, \lambda\lambda) n_{\lambda}.$$

For the ground state, i.e., for $n_0 = N$ and $n_k = 0$ ($k \neq 0$), the energy denominator is given by

$$G(\sigma, \tau) = \{ \varepsilon(0) - \varepsilon(\sigma) + \varepsilon(0) - \varepsilon(\tau) \}^{-1} \\ = - \{ k_{\sigma}^2 + \rho [u(\sigma 0, \sigma 0) + u(\sigma 0, 0\sigma) - u(00, 00)] \\ + k_{\tau}^2 + \rho [u(\tau 0, \tau 0) + u(\tau 0, 0\tau) - u(00, 00)] \}^{-1}. \quad (37)$$

It should be noted here that this Green function together with eq. (36) is just the fundamental equation for the ground state of the Bose system in Brueckner and Sawada's theory [see eqs. (2) and (3) of reference 4), p. 1128].

We have shown above that the propagator alteration is carried out by a proper choice

of the separation of the total Hamiltonian. This view-point is also useful to see how the situation is in the case of attractive interactions. By the attractive interactions we mean that the $u(00, 00)$ is negative. In this case the sound velocity becomes imaginary if one formally applies the formula for the excitation energy spectrum⁴,

$$\hbar\omega_q = (\hbar^2/2m) \{ [q^2 + \rho \{ u(0\mathbf{q}, 0\mathbf{q}) + u(0\mathbf{q}, \mathbf{q}0) - u(00, 00) \}]^2 - \rho^2 u^2(00, \mathbf{q} - \mathbf{q}) \}^{1/2}. \quad (38)$$

For in the limit $\mathbf{q} \rightarrow 0$, eq. (38) becomes

$$\hbar\omega_q = (\hbar^2/2m) q \{ 2\rho u(00, 00) \}^{1/2}.$$

Thus, the excitation energy becomes imaginary if $u(00, 00) < 0$. Such a feature is, of course, physically meaningless, hence it represents a breakdown in the approximation employed rather than a physical effect. In order to clarify this point, we first note that the assumption that nearly all the particles are in the single particle state with momentum 0 was used in deriving the excitation energy spectrum¹⁻³. This assumption may be valid for the purely repulsive potentials but is entirely misleading for the attractive potentials under consideration. This may be seen as follows.

From eq. (37) it follows that the Green function $G(\sigma, \tau)$ becomes positive for the small momentum transfer, since in the limit $\sigma, \tau \rightarrow 0$, $G(\sigma, \tau)$ approaches $-2\rho u(00, 00)^{-1}$ which is positive. Now the energy denominator represents the energy difference between the ground and excited states, and the above argument indicates that the energy value for the state in which two particles are excited is smaller than that of the ground state in which $n_0 = N$ and $n_k = 0$ ($k \neq 0$), provided that the momenta of excited particles are small. Thus, the latter state is no longer stable and does not represent the true ground state. One of the way to find this will be the variational procedure in which the energy given by eq. (33) (with v 's replaced by u 's) becomes minimum with respect to the variation of the occupation numbers. After the occupation numbers are determined, the usual procedure may be applied to the determination of excitation energy. We do not enter into this problem here, but we wish only to remark that the variation procedure in the configurational space has already been employed by various workers¹⁰.

Acknowledgments

The author wishes to express his sincere thanks to Professors S. Tomonaga and N. Fukuda for giving him a number of kind and valuable suggestions concerning this work.

References

- 1) M. Gell-Mann and K. A. Brueckner, Phys. Rev. **106** (1957), 364.
- 2) E. W. Montroll and J. C. Ward, Phys. Fluids **1** (1958), 55.
- 3) N. Bogolyubov, Jour. Exp. Theor. Phys. **34** (1958), 73.
- 4) K. A. Brueckner and K. Sawada, Phys. Rev. **106** (1957), 1117, 1128.
- 5) G. N. Watson, *Theory of Bessel Functions*, Cambridge University Press, 1922, p.402.
- 6) B. A. Lippmann and J. Schwinger, Phys. Rev. **79** (1950), 469.
K. M. Watson, Phys. Rev. **89** (1953), 575.
K. A. Brueckner and C. A. Levinson, Phys. Rev. **97** (1955), 1344.
- 7) R. Abe, Prog. Theor. Phys. **19** (1958), 713.
- 8) W. Lenz, Z. Physik **56** (1929), 778.
- 9) K. Huang and C. N. Yang, Phys. Rev. **105** (1957), 767.
- 10) R. Abe, Prog. Theor. Phys. **19** (1958), 699.
- 11) T. D. Lee, K. Huang and C. N. Yang, Phys. Rev. **106** (1957), 1135.
- 12) N. Bogolyubov, J. Phys. U. S. S. R. **11** (1947), 23.
- 13) R. Jastrow, Phys. Rev. **98** (1955), 1479.
F. Iwamoto, Prog. Theor. Phys. **19** (1958), 595.
R. Abe, Prog. Theor. Phys. **19** (1958), 57, 407.

The Clebsch-Gordan Coefficients

S. Datta MAJUMDAR

Department of Physics, University College of Science, Calcutta, India

(Received July 19, 1958)

A simpler method of deriving the Clebsch-Gordan coefficients is described. By the use of the new operators for angular momentum, introduced in a recent paper, the problem, which properly belongs to the domain of Algebra, is tackled here by the more convenient methods of Analysis. The natural occurrence of the hypergeometric function is an interesting feature of this treatment.

§ 1. Introduction

In a previous paper¹⁾ (referred to hereafter as (I)) the one-variable operators

$$M_x \pm iM_y = e^{\pm i\varphi} (j \pm iD_\varphi), \quad M_z = -iD_\varphi, \quad D_\varphi = d/d\varphi \quad (1)$$

were shown to give a representation of angular momentum both for integral and half-integral values of j^* , and it was correctly realised that their use might lead to an easier method of calculating the Clebsch-Gordan (briefly C-G) coefficients. An important step in this direction was taken by setting up a second order differential equation (eq. (4) below), whose solution would yield a general expression for these coefficients. As the very form of this equation discouraged any attempt at a solution, a recursion formula connecting three successive Fourier coefficients (which are identical with the C-G coefficients of the second kind as defined in § 2) was set up. But this formula was as intractable as the equation itself, and the possibility of getting any simplification seemed to be remote. Recently, however, it has been realised that the difficulties are illusory, and that eq. (4) is, in fact, one of the standard equations of Analysis in an apparently unrecognisable form. Furthermore, it has been possible to set up a system of $4j+2$ coupled equations, in a certain sense, more satisfactory than eq. (4). These new equations, being of the first order, can be solved immediately, and the solution involves hypergeometric functions, which naturally occur in the theory as formulated in the present paper. The entire theory can, therefore, be worked out from the known properties of these functions, and a general expression for the coefficients obtained without using the recurrence relations of Racah²⁾ or the group-theoretical method of Wigner³⁾. With the help of the transformation formulae for hypergeometric functions this expression can, in fact, be written a variety of forms, of which only one will be given here. The reduction to the hypergeometric equation places a powerful tool at our disposal and may be of help in finding out interesting relationships.

* The results of this paper are valid for half-integral quantum numbers as well.

§ 2. Clebsch-Gordan coefficients of the second kind

If the normalized spherical harmonics $Y_j^m(\theta, \varphi)$ are replaced by $e^{im\varphi}$ and the usual angular momentum operators by the operators (1), then the matrices assume the forms

$$(M_x + iM_y)_{m+1, m} = j - m, \quad (M_x - iM_y)_{m, m+1} = j + m + 1, \quad (M_z)_{mm} = m$$

where, M_x, M_y, M_z are the components of angular momentum in units of \hbar . These are connected with the usual matrices $M_x^* \pm iM_y^*, M_z^*$ by a similarity transformation $M = A^{-1}M^*A$, where A is a diagonal matrix with diagonal elements $A_{mm} = [(j-m)!(j+m)!]^{1/2}$. If $f(\varphi) = \sum_{m=-j}^j a_m e^{im\varphi}$ is the solution of the reduced problem, then the actual solution in terms of spherical harmonics is $F(\theta, \varphi) = \sum_{m=-j}^{m=j} A_{mm} a_m Y_j^m(\theta, \varphi)$ (see § 8 of (I)).

For a dynamical system with two angular momenta M_1 and M_2 the functions Ψ_{jm} , which simultaneously diagonalize $M_z = M_{z1} + M_{z2}$ and the total angular momentum $M^2 = (M_1 + M_2)^2$ are linear combinations of products of the type $Y_{j_1}^{m_1}(\theta_1, \varphi_1) Y_{j_2}^{m_2}(\theta_2, \varphi_2)$. The coefficients $\{m_1 m_2 | jm\}$ of these linear combinations are called the C-G coefficients. It will be convenient for us to call them "coefficients of the first kind." If the normalized eigenfunction Ψ_{jm} of the operator M^2 is multiplied by $[(j-m)!(j+m)!]^{1/2}$ and the replacements stated at the beginning of this section are made, we get a function

$$\Phi_{jm} = \sum_{m_1, m_2} (m_1 m_2 | jm) \exp(im_1 \varphi_1 + im_2 \varphi_2). \quad (2)$$

The coefficients $(m_1 m_2 | jm)$ of this double Fourier series will be called "C-G coefficients of the second kind." Evidently the connection between the two kinds of coefficients is

$$\{m_1 m_2 | jm\} = (m_1 m_2 | jm) \left[\frac{(j_1 - m_1)! (j_1 + m_1)! (j_2 - m_2)! (j_2 + m_2)!}{(j - m)! (j + m)!} \right]^{1/2}$$

The introduction of the coefficients of the second kind is essential for the discussions to follow and results in considerable simplification of the mathematical treatment.

§ 3. The differential equations satisfied by Φ_{jm}

We now proceed to set up certain differential equations satisfied by the function Φ_{jm} , which is an unnormalized eigenfunction of the operator

$$\begin{aligned} M^2 &= (M_1)^2 + (M_2)^2 + (M_{x1} + iM_{y1})(M_{x2} - iM_{y2}) \\ &\quad + (M_{x1} - iM_{y1})(M_{x2} + iM_{y2}) + 2M_{z1}M_{z2} \\ &= F_1 + F_2 + e^{i(\varphi_1 - \varphi_2)}(j_1 + iD_1)(j_2 - iD_2) \\ &\quad + e^{-i(\varphi_1 - \varphi_2)}(j_1 - iD_1)(j_2 + iD_2) - 2D_1D_2 \end{aligned}$$

belonging to the eigenvalue F , where,

$$F = j(j+1), \quad F_1 = j_1(j_1+1), \quad F_2 = j_2(j_2+1), \quad D_1 = \frac{\partial}{\partial \varphi_1}, \quad D_2 = \frac{\partial}{\partial \varphi_2}.$$

Introducing the variables, $\xi = \varphi_1$, $\eta = \varphi_1 - \varphi_2$, we have

$$M_x \pm iM_y = e^{\pm i\epsilon} [j_1 \pm iD_{\epsilon} \pm iD_{\eta} + e^{\mp i\eta} (j_2 \mp iD_{\eta})], \quad M_z = -iD_{\epsilon},$$

$$M^2 = F_1 + F_2 + e^{i\eta} (j_1 + iD_{\epsilon} + iD_{\eta}) (j_2 + iD_{\eta}) + e^{-i\eta} (j_1 - iD_{\epsilon} - iD_{\eta}) (j_2 - iD_{\eta})$$

$$+ 2(D_{\epsilon} + D_{\eta}) D_{\eta}.$$

The differential equations are conveniently set up with the help of the commutation relations, $[M_z, M^2] = [M_x \pm iM_y, M^2] = 0$. In combination with the equation

$$[M^2 - F]\phi_{jm} = 0 \quad (3)$$

they determine the functions ϕ_{jm} except for an arbitrary factor, which may depend on j_1, j_2, j and sometimes on m . The relation $[M_z, M^2] = -i[D_{\epsilon}, M^2] = 0$ implies that $m = m_1 + m_2$ is constant for a particular function ϕ_{jm} . The double summation over m_1 and m_2 in the defining equation (2), therefore, reduces to a simple sum, and we have

$$\phi_{jm} = e^{i\epsilon m_{\epsilon}} \sum_{m_2} (m_1 m_2 | jm) e^{-i\epsilon m_2 \eta} = e^{i\epsilon m_2 \eta} \chi_{jm}.$$

As a consequence, the variable ϵ drops out altogether from eq. (3), which reduces to

$$[e^{i\eta} (j_1 - m + iD_{\eta}) (j_2 + iD_{\eta}) + e^{-i\eta} (j_1 + m - iD_{\eta}) (j_2 - iD_{\eta})$$

$$+ 2(im + D_{\eta}) D_{\eta} + F_1 + F_2 - F] \chi_{jm} = 0. \quad (4)$$

This important equation already occurs in (I). It determines the functions χ_{jm} and therefore C-G coefficients, up to an arbitrary factor involving j_1, j_2, j, m .

Let us now examine the consequences of the pair of relations $[M_x \pm iM_y, M^2] = 0$, which imply that $(M_x \pm iM_y)\phi_{jm}$ is a linear combination of the $2j+1$ eigenfunctions ϕ_{jm} for different values of m . The occurrence of the factor $e^{\pm i\epsilon}$ in $M_x \pm iM_y$ further ensures that this linear combination consists of a single term $\phi_{m \pm 1}$. On suitably adjusting the arbitrary constant in ϕ_{jm} and dropping the variable ϵ , we have

$$[j_1 - m + iD_{\eta} + e^{-i\eta} (j_2 - iD_{\eta})] \chi_{jm} = (j - m) \chi_{jm+1} \quad (5a)$$

$$[j_1 + m + 1 - iD_{\eta} + e^{i\eta} (j_2 + iD_{\eta})] \chi_{jm+1} = (j + m + 1) \chi_{jm}. \quad (5b)$$

It is easily seen that they lead to eq. (4) and determine the functions χ_{jm} up to an arbitrary factor independent of m . This factor can be determined by normalizing the simplest of the functions ϕ_{jm} , namely, ϕ_{jj} or ϕ_{j-j} .

As it will not be necessary to consider more than one value of j at a time, it is better, at this stage, to drop the subscript j , and write χ_m for χ_{jm} , ϕ_m for ϕ_{jm} , etc. We notice that eq. (4) remains unchanged on changing the sign of m and passing on to the complex conjugate. An immediate consequence of this symmetry property is that $\chi_{-m}^* = C_m \chi_m$. Making the same changes in eq. (5a) and comparing with eq. (5b), we see that $C_m = C_{m-1} = C$ is independent of m . From $\chi_0^* = C \chi_0$ it then follows that $|C| = 1$. The exact value of C can be obtained by comparing χ_j and χ_{-j} given in eq. (9) below, and turns out to be $C = (-1)^{j_1 + j_2 - j}$. Therefore,

$$\chi_{-m} = (-1)^{j_1 + j_2 - j} \chi_m^*, \quad (-m_1 - m_2 | j, -m) = (m_1 m_2 | jm) (-1)^{j_1 + j_2 - j}. \quad (6)$$

To bring eqs. (4), (5a), (5b) into more convenient forms and to show their connection with the hypergeometric equation we put $x=e^{-\epsilon\eta}$. This gives

$$\left[\frac{1}{x}(j_1-m+xD)(j_2+xD)+x(j_1+m-xD)(j_2-xD)+2(m-xD)xD\right. \\ \left.+F_1+F_2-F\right]\chi_m=0, \quad D=d/dx, \quad (7)$$

$$L_1^m(\chi_m)=[x(1-x)D+j_1-m+j_2x]\chi_m=(j-m)\chi_{m+1}, \quad (8a)$$

$$L_2^{m+1}(\chi_{m+1})=[(1-x)D+j_1+m+1+j_2/x]\chi_{m+1}=(j+m+1)\chi_m. \quad (8b)$$

The eqs. (8a), (8b), being of the first order, can be solved easily. Putting $m=j$ in (8a) and $m=-j-1$ in (8b), we have

$$\chi_j=x^{-j_1+j}(1-x)^{j_1+j_2-j}, \quad \chi_{-j}=x^{-j_2}(1-x)^{j_1+j_2-j}. \quad (9)$$

Next, put

$$\chi_m=x^{-j_1+m}(1-x)^{j_1+j_2-j}, \quad u_m=x^{-j_2}(1-x)^{j_1+j_2-j}v_m. \quad (10)$$

This gives the following two alternative forms of the set of equations (8a), (8b):

$$\left. \begin{aligned} [(1-x)D+j-m]u_m &= (j-m)u_{m+1}, \\ [x(1-x)D-j_1+j_2+m+(j_1-j_2+j)x]v_m &= (j+m)u_{m-1}, \end{aligned} \right\} \quad (11)$$

$$\left. \begin{aligned} [x(1-x)D+j_1-j_2-m+(-j_1+j_2+j)x]v_m &= (j-m)v_{m+1}, \\ [(1-x)D+j+m]v_m &= (j+m)v_{m-1}. \end{aligned} \right\} \quad (12)$$

A comparison with the relations between contiguous hypergeometric functions, namely,

$$(xD+a)F(a, b; c; x)=aF(a+1, b; c; x),$$

$$[x(1-x)D+c-a-bx]F(a, b; c; x)=(c-a)F(a-1, b; c; x)$$

now shows that the solution of the set (11) is

$$u_m=F(-j+m, -j_1+j_2-j; -2j; 1-x)$$

and the solution of the other set (12) is

$$v_m=F(-j-m, j_1-j_2-j; -2j; 1-x).$$

We can get two other forms of the solution by making use of the relation (6). These are

$$(1-x)^{-j_1-j_2+j}\chi_m=x^{-j_1+j}F(-j+m, j_1-j_2-j; -2j; (x-1)/x) \\ =x^{-j_2-j_1+m}F(-j-m, -j_1+j_2-j; -2j; (x-1)/x).$$

The four different forms can be obtained very simply from Kummer's connection formulae.

Other forms may be obtained, for instance, from the relation

$$F(a, b; c; x)=\Gamma(c)\Gamma(c-a-b)/\{\Gamma(c-a)\Gamma(c-b)\}. \quad F(a, b; a+b-c+1; 1-x)$$

which is valid when at least one of the parameters a, b is a non-positive integer, and the

denominator in the successive terms of the series on either side of the identity does not vanish earlier than the numerator. The factor involving Gamma functions can often be evaluated by a limiting process even if it is meaningless. Each new form of the solution leads to a different expression for the C-G coefficients. But it is usually difficult to establish the equivalence of the various expressions by direct algebraic reduction.

The foregoing discussions make it clear that eq. (4) or (7) is a hypergeometric equation in disguise, and can be reduced to it by either of the substitutions (10). Elimination of χ_m or χ_{m+1} from eqs. (8a), (8b) leads to the equations

$$[L_2^{m+1}L_1^m - (j-m)(j+m+1)]\chi_m = 0, \text{ and } [L_1^{m-1}L_2^m - (j+m)(j-m+1)]\chi_m = 0$$

and these must be identical with eq. (7). The entire set of coupled first order equations can, therefore, be replaced by the single second order equation, which may be taken to be the basic equation of the problem.

§ 4. Determination of the constant A_j

As ϕ_m has a definite connection with the normalized function \mathcal{T}_m , the multiplying constant in χ_m cannot be arbitrary, but must have a definite dependence on j_1, j_2, j . To use Racah's notation we write

$$\chi_m = A_j (2j)! / (j_1 + j_2 - j)! \cdot x^{-j_1+m} (1-x)^{j_1+j_2-j} F(-j+m, -j_1+j_2-j; -2j; 1-x).$$

The constant A_j is most easily determined by normalizing the functions \mathcal{T}_j or \mathcal{T} , that is, from either of the relations

$$\sum_{m_2} \{m_1 m_2 | jj\}^2 = \sum_{m_2} \{m_1 m_2 | j, -j\}^2 = 1. \quad (13)$$

The C-G coefficient $(m_1 m_2 | j_1 j_2)$ is the coefficient of x^{m_2} in the expansion of χ_j :

$$(m_1 m_2 | jj) = A_j (2j)! (-1)^{j_1-m_1} [(j_1-m_1)! (j_2-m_2)!]. \quad (14)$$

The coefficient of the first kind is, therefore,

$$\{m_1 m_2 | jj\} = A_j (-1)^{j_1-m_1} [(2j)! (j_1+m_1)! (j_2+m_2)! / \{(j_1-m_1)! (j_2-m_2)!\}]^{1/2}.$$

The relation (13) now gives*, on writing t for $j_1 - m_1$,

$$A_j^{-2} = \sum_t \frac{(2j)! (2j_1-t)! (-j+j_2+j+t)!}{t! (j_1+j_2-j-t)!} = \frac{(2j)! (2j_1)! (-j_1+j_2+j)!}{(j_1+j_2-j)!} \times \\ \times F(-j_1-j_2+j, -j+j_2+j+1; -2j_1; 1).$$

Gauss's formula, $F(a, b; c; 1) = \Gamma(c) \Gamma(c-a-b) / \{\Gamma(c-a) \Gamma(c-b)\}$, then gives

$$A_j^{-2} = [(2j+1) (j_1+j_2-j)!]^{1/2} [(j_1-j_2+j)! (-j_1+j_2+j)! (j_1+j_2+j+1)!]^{1/2}.$$

This method of evaluation of A_j is perhaps simpler.

* Here, as in all formulae of this paper, the summation index takes all integral values consistent with the factorial notation, the factorial of a negative number being meaningless.

§ 5. General formulae for the coefficients

From the various expressions for χ_m it is easy to obtain general formulae for the C-G coefficients. In the second form they are identical with the coefficient of x^{m_2} in the expansion of χ_m . To get a formula for the coefficients we select, at random, any particular from of the solution, say,

$$\begin{aligned}\chi_m &= A_j(2j)!/(j_1+j_2-j)! \cdot x^{-j_2}(1-x)^{j_1+j_2-j} F(-j-m, j_1-j_2-j; -2j; 1-x) \\ &= \frac{A_j(-j_1+j_2+j)!(j+m)!}{(j_1+j_2-j)!} \times \\ &\times \sum_{s,t} x^{-j_2+s} (-1)^{t+s} \frac{(2j-t)!(j_1+j_2-j+t)!}{t!(j+m-t)!(-j_1+j_2+j-t)!s!(j_1+j_2-j+t-s)!}\end{aligned}$$

whence,

$$\begin{aligned}(m_1 m_2 | j m) &= \frac{A_j(-j_1+j_2+j)!(j+m)!}{(j_1+j_2-j)!(j_2+m_2)!} \\ \sum_t (-1)^{j_2+m_2+t} &\frac{(2j-t)!(j_1+j_2-j+t)!}{t!(j+m-t)!(-j_1+j_2+j-t)!(j_1-j-m_2+t)!}.\end{aligned}\quad (15)$$

The correctness of this formula can be tested by seeing if it satisfies the two recurrence relations,

$$(j-m)(m_1 m_2 + 1 | j m + 1) = (m_1 - 1 m_2 + 1 | j m)(j_1 - m_1 + 1) + (m_1 m_2 | j m)(j_2 - m_2)$$

$$(j+m)(m_1 m_2 - 1 | j m - 1) = (m_1 + 1 m_2 - 1 | j m)(j_1 + m_1 + 1) + (m_1 m_2 | j m)(j_2 + m_2),$$

which follow from eqs. (5a), (5b), and are substantially the same as Racah's relations²⁾ (3) and (5). A simple calculation shows that it satisfies the second of the two relations. Instead of trying to verify that it satisfies the other relation also it is much easier to show that it gives the correct initial conditions, that is, the expression (14) for $(m_1 m_2 | j j)$. That it does so is immediately seen by writing the expression, obtained from (15), in the form

$$\begin{aligned}(m_1 m_2 | j j) &= \frac{A_j(-j_1+j_2+j)!(2j)!}{(j_1+j_2-j)!(j_2+m_2)!} \sum_t (-1)^{j_2+m_2+t} \frac{(j_1+j_2-j+t)!}{t!(-j_1+j_2+j-t)!(j_1-j-m_2+t)!} \\ &= \frac{A_j(2j)!(-1)^{j_2+m_2}}{(j_1+j_2-j)!(j_2+m_2)!} \left[D^{j_2+m_2} \left[(1-x)^{-j_1+j_2+j} x^{j_1+j_2-j} \right] \right] x=1.\end{aligned}$$

References

- 1) S. D. Majumdar, Proc. Phys. Soc. London (in press). see also Progr. Theor. Phys. **19**, (1958), 452 (L).
- 2) G. Racah, Phys. Rev. **62** (1942), 438.
- 3) E. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*, Braunschweig (1931).

Electronic Structure of the Exciton in Ionic Crystal

—Further Discussion of a Localized Exciton (Zero-th Order Description) and
General Formulation of a Moving Exciton in Alkali Halides—

Toshinosuke MUTO, Sei'ichi OYAMA and Hiroshi OKUNO

The Institute for Solid State Physics, University of Tokyo, Tokyo

(Received July 24, 1958)

After a careful discussion of the characteristic features of the foregoing theories of an exciton in ionic crystal, a new approach to the exciton problem has been proposed, in which the one electron orbital is shown to yield already some of the characteristic properties of an exciton, an α -exciton (an exciton localized in the immediate vicinity of a negative ion vacancy) and an X-ray exciton in comparison with the experiments performed so far. In view of the present one electron orbital yielding a satisfactory zero-th order description of an exciton, a general formulation of a moving exciton based on the present one electron orbital is developed in connection with the alkali halides and an approximate estimate of a singlet exciton peak in KCl has been carried out, getting agreement with the experiment as far as our approximate evaluation of the various quantities involved is permitted. A brief discussion of the effect of spin-orbit coupling on exciton multiplet has been made and a possible interpretation is presented on the splitting into doublet of the exciton peaks in KCl and NaCl below 150°K as recently observed in the reflection method of measurement by Hartman, Nelson and Siegfried. Finally, the comparison of Slater and Shockley's theory of an exciton with the present theory is discussed in some detail and the latter is concluded to be more suitable for a small size exciton.

§ 1. Introduction

As is well known,¹⁾ two distinctly different treatments for the electronic structure of an exciton have been formulated, each with its range of applicability. An atomic orbital approach along the line first proposed by Frenkel²⁾ and recently extended by Overhauser³⁾ and others gives the solution for the lower exciton states. This treatment involves formulating the exciton wave function from atomic orbitals of the atoms at crystal lattice sites. The original Frenkel proposal considered linear combinations of one electron orbitals representing individual atoms in the crystal being excited. The formation of one electron orbitals representing an electron excited to an atomic orbital on a nearest neighbour atom (an electron and a hole on nearest neighbours) has been discussed most recently by Overhauser.³⁾ The one electron description of the latter atomic orbital approach has been dealt with already by von Hippel⁴⁾ along the classical consideration. On the other hand, an alternative treatment based on the band theory for the exciton structure was first given by Slater and Shockley.⁵⁾ Their proposal for an exciton can be most clearly pictured as an electron in a conduction band being bound to a positive hole in a valence band owing to their electrostatic attraction, which eventually leads to the effective mass approximation for the exciton problem, first given by Wannier and further discussed by Meyer,⁶⁾ Luttinger

and Kohn,⁷⁾ and Slater and Koster.⁸⁾ Recently, the extension of the effective mass approximation to the degenerate energy band edge has been discussed in detail by Dresselhaus⁹⁾ in connection with the exciton problem. The various developments in the exciton-theory, worked so far, may be considered to belong essentially to one or the other of the mentioned alternatives in the main features.

The common characteristic features of the above two methods of treatments for the exciton problem will be described in the following way. At first, the appropriate one electron orbitals are constructed in which an electron is moving in the specified electrostatic field, and then Slater's determinantal functions are set up corresponding to each electron configuration. The configuration interaction is worked in a suitable way, taking account of the effective electron configurations. The exciton characteristics in such methods of procedure do not make their appearance in the zero-order description but only in the higher order approximation through working the configuration interaction. In fact, the influence of the surrounding ions of the crystal on an excited electron has been entirely disregarded in the zero-order description of Frenkel's atomic orbital approach,²⁾ whereas, in the band theoretic approach of Slater and Shockley,³⁾ the effect of an excess positive charge is not carefully taken into account in the one electron orbitals, the field acting on an excited electron being perfectly periodic in the same way as an additive electron in ionic crystal.

Now, when we remember the fact that the self-consistent procedure of constructing the Hartree field of a crystal electron, as in the atomic structure, can hardly proceed mainly due to the mathematical difficulty but, instead, the periodic field assumption is usually adopted from the outset,¹⁰⁾ it seems to be worthwhile to investigate carefully the best one electron potential in the crystal in connection with the exciton problem. In fact, the crude consideration seems to show that, in the case of metals, an excess positive charge left after the removal of a concerned electron may be considered to be smeared out rapidly over the whole crystal owing to the screening of the remaining electrons of high mobility so that the resulting electrostatic field on a concerned electron will become sufficiently periodic in space, whereas, in the case of ionic crystals, the mentioned positive charge may be reasonably supposed to be more or less localized around an ion or a group of ions of the crystal since the electrons of ionic crystal may be regarded to be relatively strongly bound around the individual ions in the crystal. The Hartree field in ionic crystal, therefore, will probably be considered to become non-periodic in the neighbourhood of the mentioned ion or the group of ions. Thus, we shall have to deal with the perturbed periodic field for treating an electron in ionic crystal in a sense of Hartree approximation, in contrast to the periodic potential field in metallic crystal.

Our new approach to the exciton problem is to adopt the one electron orbital in such perturbed periodic field as a starting point of the conventional method of treatment of many electron system in ionic crystal. In fact, a suitably chosen perturbed periodic field for potassium chloride has been shown, in Part I,¹¹⁾ to give rise to a number of localized states just below the conduction band, which have actually some of the exciton characteristics observed so far and are found to correspond to Dexter's excitation model.¹²⁾

The comparison with experiment seems to show good agreement under a fewer assumptions compared with the case of von Hippel's model.⁽¹⁰⁾ In order to extend the validity of our method of procedure for working the zero-order description of an exciton in alkali halides, the similar calculations to the previous one are developed to all alkali chlorides and, further, to an exciton localized around a negative ion vacancy in alkali chlorides which corresponds to one of the Greek bands observed by Pringsheim and others⁽¹¹⁾ in alkali iodides and bromides; their results shall be described in the next chapter together with some remarks on the X-ray exciton in KCl.⁽¹²⁾ As a next higher approximation, we shall deal with a moving exciton by taking into account the configuration interaction based upon the one electron orbital mentioned above. The actual procedure of treating such moving exciton will be carried through in a straightforward way and leads to the characteristic secular equation from which the exciton multiplet can be deduced, at least in principle, in view of the branches of a valence band and the spin-orbit coupling. The result seems to roughly correspond to some refinement of the pioneer work of Frenkel and will be reasonable particularly for the small size exciton or the lower exciton states in the crystal of relatively small dielectric constant.

In this connection, it should be mentioned that the exciton problem seems to be worthwhile paying attention for the following reasons: first, it can provide us very important features of crystal electrons which can not be satisfactorily dealt with by the simple band theory which has been fruitful for solid state physics, and secondly it is regarded as one of the typical problems of the correlation effect of crystal electrons which can be manipulated mathematically through taking into account the configuration interaction, although a strong correlation in large scale of crystal electrons is shown to be responsible for the collective oscillation of plasma type as shown in Bohm and Pines' theory.⁽¹³⁾

§ 2. Localized exciton in alkali halides

(Zero-order description of an exciton)

a) *Optical exciton in alkali chlorides.**

According to the consideration described in Introduction, our method of procedure in Part I⁽¹⁾ has been extended to all alkali chlorides which have actually been observed experimentally. Our main task is to solve the eigenvalue problem given by Wannier-Slater's equation valid for alkali chlorides which involves, as parameters, effective mass m^* near a conduction band bottom and effective dielectric constant κ responsible for the polarization of surrounding medium. Our numerical results are shown in Table I in which the observed values are quoted from Hilsch and Pohl, and Schneider and O'Bryan.⁽¹⁴⁾ $-E'$ represents the energy difference between a ground state of an exciton and a conduction band bottom. According to the recent experiment by Hartman et al.,⁽¹⁵⁾ however,

* The brief account of the present section has been reported in Short Notes of J. Phys. Soc. Japan, 12 (1957), 108.

the absorption peak corresponding to a series limit of exciton states in NaCl seems to vanish, which fact requires further careful considerations from both experimental and theoretical sides.

Table I

	κ	m^*/m	$-E_{\text{cul.}}$	$-E_{\text{obs.}}$
LiCl	2.75	0.6	0.89 eV	0.86 eV
NaCl	2.25	1.0	1.81	1.83
KCl	2.13	0.85	1.83	1.80
RbCl	2.19	0.7	1.53	1.50
CsCl	2.60	0.9	1.33	1.39

At first sight, it might be argued that a variety of the possible values for two parameters m^* and κ may happen to give equally good fit to the experimental results. However, a general consideration, both experimental and theoretical, requires for κ to lie in the range of one to high frequency values and for m^* to be of the order of magnitude ranging from $0.5m$ to m , m being the true mass of an electron, which situation enables one to get almost unique assignment of these quantities in our method of procedure. According to the previous experience¹¹⁾ in the case of KCl, the effective dielectric constants κ have been assumed to be high frequency values in our results shown in Table I. The assigned values of m^* are observed to be of the order of magnitude to be expected in accordance with the available other experimental evidences. Generally speaking, the comparison shown in the table seems to be good for the zero-order description of an exciton as far as the approximation in our method of procedure is allowed for.

b) X-ray exciton in KCl.

In Part I we have worked the numerical estimates of the location of a ground state of X-ray exciton from the bottom of a conduction continuum in KCl in comparison with those of Parratt and Jossem's observation.¹⁴⁾ The results were as follows.

	Our estimate	Parratt-Jossem's one
K ⁺	3.39 eV	3.2 eV
Cl ⁻	2.5 eV	4.13 eV

Our estimates were obtained by tentatively assuming $0.8m$ for the effective mass near the conduction continuum bottom and 1.5 for the effective dielectric constant in view of the short period of X-rays. The discrepancy seems to appear in the case of Cl⁻. However, since the bottom position of a conduction continuum is not entirely recognizable in the observed contour for either chlorine or potassium ions of KCl, Parratt and Jossem seem to have actually made use of the following empirical relation of the width of a valence band (ΔE_v) with the location of an X-ray exciton peak due to Cl⁻ ion (A peak in their contour) from the bottom of a conduction continuum ($|E|'$), which relation is

established through the careful analysis of both K_α emission and K absorption spectra in Cl ion of KCl together with its ultraviolet absorption spectra.

$$9.46 - |E'| + \frac{1}{2}E_v = 9.75 \text{ (in eV)},$$

in which 9.46 eV represents the energy gap between the conduction and valence bands, estimated from the second peak of the ultraviolet absorption and 9.75 eV the energy distance from the valence band bottom to a X-ray exciton around Cl ion (A peak), estimated from K_α emission line and A peak of X-ray absorption in Cl^- . If the above empirical relation were accepted, we would have very interesting checking for the location of the conduction continuum bottom, which depends on the width of a Cl 3p valence band in KCl through the above relation. In fact, Parratt and Jossem seem to have supposed for the width of a Cl 3p band to be 4.42 eV, probably suggested by Shockley's computation of the energy band width in NaCl, which resulted into 4.13 eV for the location of a conduction band bottom from the exciton peak, as mentioned above. On the other hand, the substitution of our estimate of E' , 2.5 eV, in the above relation leads to the width of valence band of 2.79 eV which is narrower than the above mentioned one. When, as suggested by Parratt and Jossem,²² the polarization around a chlorine ion lacking a 1s electron is taken to become larger compared with that around a potassium ion lacking a 1s electron, owing to the larger volume of the former, and provisionally we take $\kappa=2.0$ and $m^*=0.8m$ for the former, the interpolation from Table I of Part I leads to the X-ray exciton state at about 1.7 eV below the bottom of a conduction continuum and, consequently, to the valence band width of about 2 eV in KCl by using the empirical relation. On the other hand, a recent theoretical calculation by Howland²³ of the energy bands in KCl by the use of the Bloch sum method indicates increasing favour for a narrow valence band of KCl, i.e., 1.52 eV for the width. In view of the above situation, our estimate for the location of X-ray exciton relative to the continuum bottom seems to be on a right line, though qualitatively, as far as the zero-order description of an exciton is concerned.

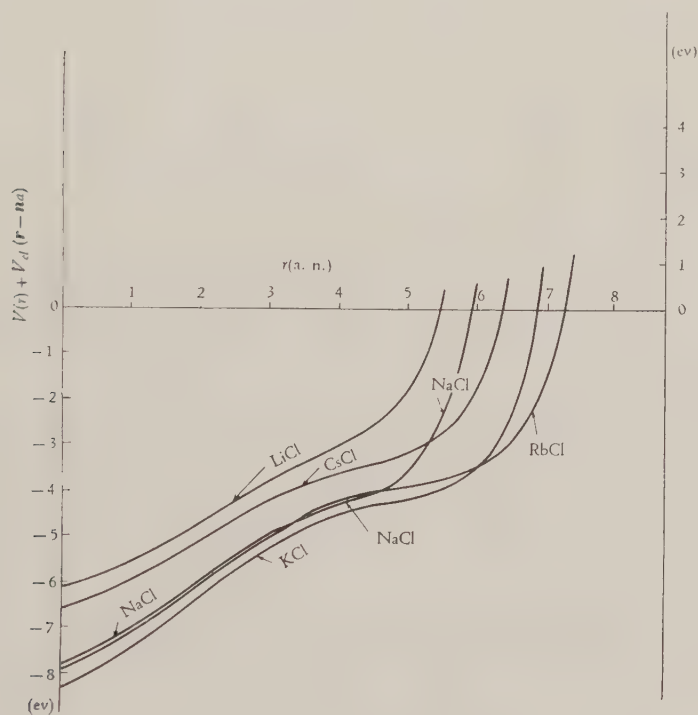
c) *An exciton localized in the immediate vicinity of a negative ion vacancy in alkali chlorides (α exciton).**

Recently, Bassani and Inchausti²⁴ have calculated the energies to create an exciton localized in the immediate vicinity of a negative ion vacancy or an F -center by a similar cycle method as was adopted in the semi-classical treatment of von Hippel. In fact, the appearance of the new absorption bands in the tail part of the fundamental absorption when a large number of halogen vacancies or F -centers are present in the crystal seems to be ascribed to such localized excitons. In view of the above situation, our method of procedure mentioned above has been applied to such case. For an exciton trapped in the immediate vicinity of a negative ion vacancy, the potential field in the Wannier-Slater equation is taken to be the negative of the Hartree field²⁵ due to the nearest neighbouring

* The brief account of this section was read at the annual meeting of Physical Society of Japan, October 17, 1957.

Cl ion, i. e., $V_{cl}(\mathbf{r}-\mathbf{na})$ in addition to the potential field of a normal exciton which has been worked out in Part I, i. e., $V(r)$. Furthermore, the former potential is assumed to be reduced by a static dielectric constant of the crystal, while, the latter one by an effective dielectric constant as was done in Part I. The resulting potential is easily seen to extend toward a neighbouring negative ion vacancy, as described in Fig. 1.

Fig. 1



The Wannier-Slater equation for the present case becomes

$$-\frac{\hbar^2}{2m^*} \Delta \psi(\mathbf{r}) + \{V(r) + V_{cl}(\mathbf{r}-\mathbf{na})\} \psi(\mathbf{r}) = E' \psi(\mathbf{r}), \quad (1)$$

whose solution has been obtained, for the ground state, by the variational principle. The trial function is assumed to be given by a superposition of a p -type function on an s -type one, taking into account the fact that the attractive field extends toward a negative ion vacancy. The obtained numerical results are shown, for alkali chlorides, in Table II, together with the results of Bassani and Inchauspe (BI) and the observed values (JE_{exp}). E' (α -exciton) and E' (free exciton) represent the energies of an α -exciton and a normal exciton from the conduction band bottom, and $JE' = |E'(\alpha\text{-exciton}) - E'(\text{free exciton})|$. The comparison seems to be rather good, allowing for the approximate nature of our method of procedure. Unfortunately, however, Greek bands in alkali chlorides have not yet been reported, probably because they lie in the vacuum ultra-violet. Finally, it should

Table II

	E' (α -exciton)	E' (free exciton)	$\Delta E'$	BI	ΔE_{exp}
LiCl	-1.19 ev	-0.89 ev	0.30 ev	—	—
NaCl	-2.39	-1.81	0.58	0.50 ev	—
KCl	-2.37	-1.83	0.54	0.53	—
RbCl	-2.09	-1.53	0.56	—	—
CsCl	-1.76	-1.33	0.23	—	—
KBr	—	—	—	0.49	0.59
KI	—	—	—	0.44	0.59

be mentioned that the rather rapidly varying region of the potential in (1) has been found to be less sensitive to the energy determination and then the approximate validity of the Wannier-Slater equation seems to hold.

Summing up the results obtained so far, the one electron orbital in a suitably constructed, perturbed periodic field may be considered to afford a good zero-order description of an exciton in alkali halides, with fewer assumptions and much smaller corrections than the foregoing models. Furthermore, according to the recent discussion by Dexter,¹² the oscillator strength and charge distribution of the present model have been shown to become a satisfactory zero-order description of an exciton. In the present calculations, however, our main concern lies in the electronic structure of an exciton in alkali chlorides in which a chlorine ion has relatively small spin-orbit interaction. In order to work the exciton in alkali bromides or iodides which show sizable fine structures, we shall have to take into account the spin-orbit interaction and spin-spin interaction between an electron and a hole within a localized exciton, by introducing them appropriately into our Wannier-Slater equation. However, such higher order interactions should be taken into consideration in more reasonable way in a many electron treatment of the following chapter.

§ 3. A moving exciton

a) General consideration.*

As stressed in the previous chapter, a localized exciton discussed above is regarded as the zero-order description of an exciton. However, a localized exciton state in unit cell is energetically equivalent to that in any other unit cell as a result of translation symmetry of the crystal lattice, so that such a state of excitation will propagate through the crystal. Next we shall proceed to work a mathematical formulation of such a moving exciton in the crystal on the basis of the one electron orbital described above.

In the following computation of a moving exciton, therefore, the wave function of the localized exciton state just below the conduction band is adopted for one electron orbitals of the excited states. The normal states of valence electrons are supposed to be

* The brief account of this section has been reported as Short Notes in J. Phys. Soc. Japan, 12 (1957), 108.

approximately described by the Bloch orbitals of the simple band theory since the majority of valence electrons of the crystal is expected to lie in the valence continuum which behaves quite similarly to the valence band of simple band theory, and a few of them is brought into the localized states just below the valence continuum, as revealed in the K_α emission line¹⁴⁾ in KCl. Furthermore, the inner shell electrons within each ion of the crystal will be amalgamated into the atomic cores whose interaction with valence electrons is represented by a suitable potential field $V_p(\mathbf{r})$.

Then, Schrodinger wave equation of N valence electrons in the crystal may be written as

$$\mathbf{H}\Psi = \left\{ -\frac{\hbar^2}{2m} \sum_{i=1}^N \Delta_i + \sum_{i=1}^N V_p(\mathbf{r}_i) + \sum_{i>j}^N \frac{e^2}{r_{ij}} \right\} \Psi = E\Psi, \quad (2)$$

where i specifies each of the valence electrons, $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ and the other quantities have the usual meanings.

As for the available electron configurations, we shall confine ourselves to the configurations with one electron excited to a ground state of the localized exciton below the conduction band from any one level of the valence band. Different configurations will be specified by different hole states of the valence band and by different locations of the localized exciton. The other possible configurations in which an electron is raised into the various excited states of a localized exciton and the conduction continuum, and, further, two or more electrons are raised into the similar excited states, will be entirely disregarded mainly for the sake of mathematical simplicity. However, the generalization of the present formalism by allowing for the neglected configurations will be easily carried out and will lead to the secular equation of the higher degree. The mentioned, approximate procedure of a single exciton will be safely applicable to the lowest exciton bands as far as the concentration of existing excitons is so low that their mutual interactions need not be taken into account. Throughout the remainder of the discussion we shall refer to KCl, regarding it as a prototype of the other alkali halides.

As actually shown in Part I,¹¹⁾ the one electron orbital of a localized exciton around \mathbf{R}_l becomes

$$c(\mathbf{l}; \mathbf{r}) = \sum_l U(\mathbf{R}_l - \mathbf{R}_i) \cdot a_c(\mathbf{r} - \mathbf{R}_i) \quad (3)$$

which satisfies the Schrodinger equation of the perturbed periodic lattice,

$$\left\{ -\frac{\hbar^2}{2m} \Delta + V_p'(\mathbf{r}) + \Delta V(\mathbf{r} - \mathbf{R}_l) \right\} \cdot c(\mathbf{l}; \mathbf{r}) = E^{\text{exc}} c(\mathbf{l}; \mathbf{r}) \quad (4)$$

and the Bloch orbital of the n -th branch of a valence band is given by

$$b_n(\mathbf{k}; \mathbf{r}) = G^{-3/2} \sum_l e^{i\mathbf{k} \cdot \mathbf{R}_l} a_n(\mathbf{r} - \mathbf{R}_l) \quad (5)$$

which satisfies

$$\left\{ -\frac{\hbar^2}{2m} \Delta + V_p'(\mathbf{r}) \right\} b_n(\mathbf{k}; \mathbf{r}) = E_n(\mathbf{k}) b_n(\mathbf{k}; \mathbf{r}) \quad (6)$$

since a Cl 3*p* band, i. e., valence band has been shown to split into three branches by Bethe⁽²²⁾ and Howland⁽¹⁰⁾ using the Bloch sum method, each branch of the valence band being designated by a suffix *n*. $V_n'(\mathbf{r})$ expresses the periodic potential of an additive electron in KCl and $UV(\mathbf{r}-\mathbf{R}_l)$ the potential due to a positive hole localized around a lattice point \mathbf{R}_l , which has been actually worked out in Part I. $a(\mathbf{r}-\mathbf{R})$ and $a_n(\mathbf{r}-\mathbf{R}_l)$ represent Wannier functions around \mathbf{R} , belonging to a conduction band and the *n*-th branch of a valence band, respectively. $U(\mathbf{R}_l-\mathbf{R}_l)$ has been shown, in Part I, to become of $\exp(-Z|\mathbf{R}_l-\mathbf{R}_l|a_H)$ for the ground state of a localized exciton. Furthermore, E^{exc} and $E_n(k)$ express the energies of a ground state of a localized exciton and of the *n*-th branch of a valence band, respectively, *k* being wave number vector. It should be mentioned that one electron orbital of a localized exciton (3) is orthogonal to the Bloch orbital of a valence band (5), as clearly shown by the orthogonality property of the corresponding Wannier functions.

For the ground state of the crystal as a whole, the valence band is completely filled with electrons and a localized exciton state or a conduction continuum is entirely empty. The corresponding electron configuration corresponds to a singlet state, of which the determinantal functions in the Bloch orbital method and in the Wannier function method are identical with each other.

The corresponding energies become

$$E_g = \sum_{n, \mathbf{R}_l} (a_n(\mathbf{r}_i - \mathbf{R}_l), H(i)a_n(\mathbf{r}_i - \mathbf{R}_l)) \\ + \sum_{n, \mathbf{R}_l} \sum_{m, \mathbf{R}_l} \left\{ (a_n(\mathbf{r}_1 - \mathbf{R}_l), a_m(\mathbf{r}_2 - \mathbf{R}_l), e^2/r_{12}a_n(\mathbf{r}_1 - \mathbf{R}_l)a_m(\mathbf{r}_2 - \mathbf{R}_l)) \right. \\ \left. - (a_n(\mathbf{r}_1 - \mathbf{R}_l), a_m(\mathbf{r}_2 - \mathbf{R}_l), e^2/r_{12}a_n(\mathbf{r}_2 - \mathbf{R}_l)a_m(\mathbf{r}_1 - \mathbf{R}_l)) \right\}, \quad (7)$$

in which *n* or *m* specifies each branch of a valence band as well as the corresponding spin state, and $H(i) = -\hbar^2/2m \cdot \Delta_i + V_p(\mathbf{r}_i)$.

For the electron configurations corresponding to the excited states, let us suppose that an electron is excited from a *k_i* state of the *n*-th branch of a valence band to a localized exciton level around a lattice point \mathbf{R} . Then, according to the different spin orientations of both an excited electron and a positive hole, we can construct as usual the determinantal functions belonging to singlet and triplet states, which consist of the Bloch orbitals of a valence band and a localized exciton wave function. When we multiply each determinantal function mentioned above by $\exp(i\mathbf{k}_i \cdot \mathbf{R}_l)$ and then sum up over *k_i*, the resulting functions are shown, using Slater and Shockley's procedure,⁸ to be proportional to the determinantal functions built up from the Wannier functions themselves, belonging to a valence band, i. e., $a_n(\mathbf{r}-\mathbf{R}_l)$ and a localized exciton wave function, $c(l, \mathbf{r})$. The latter determinantal functions contain an excited electron in a localized exciton state around \mathbf{R} , and a positive hole in the state represented by a Wannier function around \mathbf{R}_l of the *n*-th branch of a valence band. Thus, the determinantal functions corresponding to singlet and triplet states of the crystal as a whole may be written as

$${}^1\mathcal{P}^{(nl,el)} = \frac{1}{\sqrt{2}} \{ \mathcal{Q}^{(nl\downarrow; el\downarrow)} + \mathcal{Q}^{(nl\uparrow; el\uparrow)} \} \quad (8)$$

and

$${}^3\mathcal{P}_0^{(nl,el)} = \frac{1}{\sqrt{2}} \{ \mathcal{Q}^{(nl\downarrow; el\downarrow)} - \mathcal{Q}^{(nl\uparrow; el\uparrow)} \}, \quad (9)$$

$${}^3\mathcal{P}_{+1}^{(nl,el)} = \mathcal{Q}^{(nl\downarrow; el\uparrow)}, \quad (9)'$$

$${}^3\mathcal{P}_{-1}^{(nl,el)} = \mathcal{Q}^{(nl\uparrow; el\downarrow)}, \quad (9)''$$

respectively. $\mathcal{Q}^{(nl\uparrow; el\downarrow)}$ represents the determinantal function built up from the Wannier functions and a localized exciton wave function, by assuming an electron with down-spin to be removed from the state represented by $a_n(\mathbf{r}-\mathbf{R}_l)\beta(\sigma)$ and to be present in a ground state with down-spin of a localized exciton, i.e., $c(\mathbf{l}; \mathbf{r})\beta(\sigma)$. $\alpha(\sigma)$ and $\beta(\sigma)$ represent spin functions and $(nl\downarrow)$ specifies a hole state and $(el\downarrow)$ an excited electron state, respectively. Its explicit expression becomes

$$\begin{aligned} \mathcal{Q}^{(nl\downarrow; el\downarrow)} = & \frac{1}{\sqrt{N!}} \sum_p \delta_p \cdot p \left\{ a_1(\mathbf{r}_1-\mathbf{R}_1)\alpha(\sigma_1) \cdots a_n(\mathbf{r}_{p-1}-\mathbf{R}_l)\alpha(\sigma_{p-1}) \right. \\ & \left. \times c(\mathbf{l}; \mathbf{r}_p)\beta(\sigma_p) a_n(\mathbf{r}_{p+1}-\mathbf{R}_{l+1})\alpha(\sigma_{p+1}) \cdots a_3(\mathbf{r}_N-\mathbf{R}_{N_0})\beta(\sigma_N) \right\} \end{aligned} \quad (10)$$

in which $a_n(\mathbf{r}_p-\mathbf{R}_l)\beta(\sigma_p)$ is replaced by $c(\mathbf{l}; \mathbf{r}_p)\beta(\sigma_p)$ in the ground state function. The other determinantal functions in (8) and (9) are similarly defined, corresponding to the various spin orientations of a positive hole in a state represented by the Wannier function of a valence band and an excited electron in a localized exciton state.

Allowing for the configuration interaction, we can set up the following linear combinations of (8) and (9) in order to solve the Schroedinger wave equation (2).

$${}^i\mathcal{P} = \sum_{nl} {}^iA_{nl} \cdot {}^i\mathcal{P}^{(nl,el)}, \quad i=1, 3 \quad (11)$$

which takes into account the interaction among the various configurations specified by the different locations \mathbf{R}_l of a localized exciton and a positive hole, and by the different branches n of a valence band in which a positive hole is situated.

Substitution of (11) in (2) leads, as usual, to the following secular equation.

$$\sum_{n'l'} ({}^i\mathcal{P}^{(nl,el)}, \mathbf{H} {}^i\mathcal{P}^{(n'l',el')}) {}^iA_{n'l'} = {}^iE {}^iA_{nl}. \quad (12)$$

In view of the translation symmetry of the energy matrix elements in (12), we can put, as usual,

$${}^iA_{nl} = e^{i\mathbf{K}\cdot\mathbf{R}_l} {}^iA_n \quad (13)$$

in which \mathbf{K} represents a translation vector in the reciprocal lattice or a wave number vector of a moving exciton.

Then, it follows

$$\sum_{n'} {}^iA_{n'} \sum_{l'} ({}^i\mathcal{P}^{(nl,el)}, \mathbf{H} {}^i\mathcal{P}^{(n'l',el')}) e^{i\mathbf{K}\cdot(\mathbf{R}_{l'}-\mathbf{R}_l)} = {}^iA_n {}^iE \quad (14)$$

which leads to a 3×3 secular determinant,

$$\det \left| \sum_l ({}^i\mathcal{P}^{(nl, el)}, \mathbf{H} {}^i\mathcal{P}^{(nl'l, el')}) e^{i\mathbf{K} \cdot (\mathbf{R}_{ll} - \mathbf{R}_l)} - {}^iE \cdot \delta_{nl, n'l} \right| = 0. \quad (15)$$

The diagonal elements of the energy matrix becomes, after some process of calculations,

$$\begin{aligned} E^{(n, n')} &= ({}^i\mathcal{P}^{(nl, el)}, \mathbf{H} {}^i\mathcal{P}^{(nl, el)}) + \sum_{l' (\neq l)} e^{i\mathbf{K} \cdot (\mathbf{R}_{ll'} - \mathbf{R}_l)} ({}^i\mathcal{P}^{(nl, el)}, \mathbf{H} {}^i\mathcal{P}^{(nl'l, el')}) \\ &= E_g + E^{\text{exc}} - \varepsilon_n(0) \pm (el, nl | e^2/r_{12} | nl, el) + (\partial V_{el} - \partial V_h) \\ &\quad + \sum_{l' (\neq l)} e^{i\mathbf{K} \cdot (\mathbf{R}_{ll'} - \mathbf{R}_l)} [-S_{ll'} \cdot \varepsilon_n(\mathbf{R}_{ll'} - \mathbf{R}_l) - S_{ll'} \sum_{r (\neq nl'l\alpha)} \{ (r, nl' | e^2/r_{12} | r, nl) \\ &\quad - (r, nl' | e^2/r_{12} | nl, r) \} + \frac{2}{0} (el, nl' | e^2/r_{12} | nl, el') - (el, nl' | e^2/r_{12} | el', nl)], \end{aligned} \quad (16)$$

where 2 or 0 and \pm signs stand for singlet and triplet exciton states, respectively.

$$\varepsilon_n(\mathbf{R}_j) = G^{-3} \sum_{\mathbf{k}} E_n(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{R}_j},$$

$$S_{ll'} = \int c^*(l'; \mathbf{r}) c(l; \mathbf{r}) d\mathbf{r}, \quad (\text{Overlap-integral}),$$

$$(el, nl | e^2/r_{12} | nl, el) = \int c^*(l; \mathbf{r}_1) a_n^*(\mathbf{r}_2 - \mathbf{R}_l) e^2/r_{12} \cdot c(l; \mathbf{r}_2) a_n(\mathbf{r}_1 - \mathbf{R}_l) d\mathbf{r}_1 \cdot d\mathbf{r}_2, \quad (17)$$

$$(el, nl' | e^2/r_{12} | nl, el') = \int c^*(l; \mathbf{r}_1) a_n^*(\mathbf{r}_2 - \mathbf{R}_l) e^2/r_{12} \cdot c(l'; \mathbf{r}_2) a_n(\mathbf{r}_1 - \mathbf{R}_l) d\mathbf{r}_1 \cdot d\mathbf{r}_2,$$

etc.

The summation with respect to r in (16) extends over all states including spin except for $a_n(\mathbf{r} - \mathbf{R}_l) \alpha(\sigma)$ and, furthermore, we have

$$\partial V_{el} = (el | V_p - V_p' | el) + \sum_{r (\neq nl'l\alpha)} \{ (r, el | e^2/r_{12} | r, el) - (r, el | e^2/r_{12} | el, r) \} \quad (18)$$

and

$$\partial V_h = (nl | V_p - V_p' | nl) + \sum_{r} \{ (r, nl | e^2/r_{12} | r, nl) - (r, nl | e^2/r_{12} | nl, r) \}. \quad (19)$$

On the other hand, the off-diagonal elements of the energy matrix may be written

$$\begin{aligned} E^{(n, n')} &= - \sum_{r (\neq nl'l\alpha)} \{ (r, n'l | e^2/r_{12} | r, nl) - (r, n'l | e^2/r_{12} | nl, r) \} \\ &\quad + \frac{2}{0} (el, n'l | e^2/r_{12} | nl, el) - (el, n'l | e^2/r_{12} | el, nl) \\ &\quad + \sum_{l' (\neq l)} e^{i\mathbf{K} \cdot (\mathbf{R}_{ll'} - \mathbf{R}_l)} [-S_{ll'} \sum_{r (\neq nl'l'l\alpha)} \{ (r, n'l' | e^2/r_{12} | r, nl) - (r, n'l' | e^2/r_{12} | nl, r) \} \\ &\quad + \frac{2}{0} (el, n'l' | e^2/r_{12} | nl, el') - (el, n'l' | e^2/r_{12} | el', nl)]. \end{aligned} \quad (20)$$

According to the valence band formula⁽²²⁾ obtained by the Bloch sum method, $\varepsilon_n(0)$ represents the sum of the atomic energy of a $3p$ electron within a chlorine ion and the Coulomb energy of the corresponding electron in the field of all negative and positive

ions in the crystal except for a chlorine ion to which the electron belongs. Such Coulomb energy will be approximated by Madelung's potential V_M at a lattice site of a chlorine ion. Then, we have

$$\varepsilon_n(0) \simeq E_{\text{atom}}(3p) + V_M, \quad V_M = -\alpha e^2/r_0, \quad (21)$$

where $\alpha=1.784$ for NaCl type lattice and r_0 denotes the cation-anion distance.

The first term on the right hand side of (18) represents the negative of the Coulomb and exchange energies of an electron within a localized exciton around R_i in the charge distribution of all the remaining valence electrons, whereas, the second term of it expresses the similar potential energy computed by the Wannier functions belonging to a valence band. Similarly, the first and second terms on the right hand side of (19) represent the corresponding quantities of a positive hole in the n -th branch in the Hartree-Fock charge distribution of all the valence electrons. If the Coulomb and exchange energies in the second terms of (18) and (19) were evaluated by using the Hartree-Fock solutions instead of Wannier functions, we would have exactly $\partial V_{el} = \partial V_h = 0$. Otherwise, ∂V_{el} and ∂V_h will not necessarily vanish but will be expected to remain small in our approximation. Therefore, $(\partial V_{el} - \partial V_h)$ may be reasonably supposed to become still smaller, on account of which we shall neglect it approximately in the following discussion. Each term of the first line of (16) has fairly simple interpretation on the atomic viewpoint, of which the first term represents the ground state energy of N electrons, the second the one electron energy of a ground state of localized exciton, the third the energy of a positive hole in the $3p$ state of a chlorine ion within the crystal and the fourth the exchange energy between an excited electron and a positive hole around the same lattice site. The remaining expressions of (16) represent the interaction energies due to the electrons in the various states with different locations in the crystal. Further, in (20) are involved the Coulomb and exchange interaction energies among the states belonging to different branches of a valence band and to a localized exciton.

Generally speaking, the 3×3 secular equation will have at most three roots for the energy of a moving exciton, corresponding to each value of \mathbf{K} . Thus, the three branches of exciton band could be expected to appear as far as our approximate configuration interaction is valid. However, when we take into account the effect of spin-orbit coupling and, further, the contribution from the higher order configurations in which an electron is raised into the excited states of a localized exciton and a conduction continuum from any one state of the valence band, the corresponding secular equation is readily seen to become of the higher order, which results into an increasing number of splitting of the exciton band.

In order to solve numerically the secular equation for the exciton band, we are actually required to work out the Wannier functions belonging to both valence and conduction bands. One of the authors* has recently worked out the Wannier function of a valence band KCl, by making use of the Bloch orbital obtained by Howland.¹⁰⁾ His Wannier function of localized type may be written as

* S. Oyama: *Busseiron Kenkyu* (1958) (in press).

$$\begin{aligned}
a_1(\mathbf{r}) = & 0.8 \cdot \phi_x(0) + 0.086 \cdot (\varphi_x(100) + \varphi_x(\bar{1}00)) \\
& - 0.016 \cdot (\varphi_x(010) + \varphi_x(0\bar{1}0) + \varphi_x(001) + \varphi_x(00\bar{1})) \\
& + 0.003 \cdot (\phi_x(110) + \phi_x(\bar{1}\bar{1}0) + \phi_x(\bar{1}\bar{1}0)) \\
& + \phi_x(1\bar{1}0) + \phi_x(101) + \phi_x(10\bar{1}) + \phi_x(\bar{1}0\bar{1}) + \phi_x(\bar{1}01)) \\
& - 0.008 \cdot (\phi_x(011) + \phi_x(0\bar{1}1) + \phi_x(0\bar{1}\bar{1}) + \phi_x(01\bar{1})) \\
& - 0.089 \cdot (\varphi_x(111) + \varphi_x(\bar{1}\bar{1}1) + \varphi_x(1\bar{1}1) + \varphi_x(11\bar{1})) \\
& + \varphi_x(\bar{1}\bar{1}1) + \varphi_x(1\bar{1}\bar{1}) + \varphi_x(\bar{1}\bar{1}\bar{1}) + \varphi_x(\bar{1}\bar{1}\bar{1})) \\
& + 0.054 \cdot (\phi_x(200) + \phi_x(\bar{2}00) + 0.023 (\phi_x(020) + \phi_x(0\bar{2}0) + \phi_x(002) + \phi_x(00\bar{2}))) \\
& + \text{higher order terms involving } \phi_y, \phi_z, \varphi_y \text{ and } \varphi_z,
\end{aligned} \tag{22}$$

in which $\phi_x(l, m, n)$ represents a $3p$ atomic orbital of a chlorine ion at a lattice point (l, m, n) , i. e., $\phi_{3p_x}(\mathbf{r} - \mathbf{R}_{lmi})$ and similarly $\varphi_x(l, m, n)$, a $3p$ atomic orbital of a potassium ion at a lattice point (l, m, n) . $a_2(\mathbf{r})$ and $a_3(\mathbf{r})$ are obtained from the above equation through a cyclic exchange of x , y and z . Thus, the approximate form of Wannier functions will be taken as

$$\left. \begin{aligned} a_1(\mathbf{r}) &\simeq 0.8 \cdot \phi_x(0) \\ a_2(\mathbf{r}) &\simeq 0.8 \cdot \phi_y(0) \\ a_3(\mathbf{r}) &\simeq 0.8 \cdot \phi_z(0) \end{aligned} \right\}. \tag{23}$$

Now, the absorption coefficient in the exciton region can be readily computed using (11) and (13), with the following result.

$$\frac{c}{4\pi} \eta \propto \left| \sum_{n=1}^3 A_n \int a_n^*(\mathbf{r}) \cdot \text{grad } c(0; \mathbf{r}) d\mathbf{r} \right|^2, \tag{24}$$

where η denotes the absorption coefficient, c the light velocity, \mathbf{q} the wave number vector of an incident light and the conservation laws of both energy and wave number vector are given by

$$h\nu_0 = {}^1E(\mathbf{k}=0) - E_g, \quad \mathbf{k} = 2\pi\mathbf{q} \simeq 0. \tag{25}$$

In view of the above situation, we shall discuss in more detail the exciton energy corresponding to $\mathbf{K}=0$, which is responsible for the peak frequency in the case of neglecting the interaction with lattice vibrations. Using (23) for the Wannier functions of a valence band and assuming for those of a conduction band to be approximately of an s -type symmetry around each lattice point, we can easily see that the matrix elements of e^2/r_{12} with respect to the Wannier functions belonging to different branches of a valence band and a localized exciton wave function with the same location are to vanish from the symmetry consideration. Furthermore, the summation over \mathbf{l} or \mathbf{R}_i of the similar matrix elements with respect to the Wannier functions or exciton wave functions of the different locations is seen to tend to vanish owing to both cubic-symmetrical arrangement

of ions in the crystal and the p -type character of (23), a pair of matrix elements symmetrically located being canceled with each other. Generally speaking, it will be permissible, for the qualitative discussion, to disregard the minor contributions from the above mentioned quantities in (16) and (20).

Then, it follows, with neglect of the small quantities involved,

$$E^{(n,n)}(\mathbf{K}=0) \simeq E_g + E^{\text{exc}} - (E_{\text{atom}}(3p) + V_M) \pm (el, nl|e^2/r_{12}|nl, el) - \sum_{l'(\neq l)} S_{ll'} \cdot \varepsilon_n(\mathbf{R}_{ll'} - \mathbf{R}_l) + \mathcal{J}E \quad (26)$$

and

$$E^{(n,n)}(\mathbf{K}=0) \simeq 0, \quad (27)$$

which shows that the singlet exciton states corresponding to $\mathbf{K}=0$ are nearly degenerate and the peak frequency does not split when we disregard the effect of spin-orbit coupling and the interaction with lattice vibrations. The triplet exciton state lies lower than the singlet one, by an amount of twice the exchange energy of an excited electron and a positive hole. As far as the main terms (the second and the third ones) in the energy expression (26) is concerned, the transition to an exciton state $(E^{(n,n)}(\mathbf{k}=0) - E_g)$ may be understood approximately as a transition of an electron from a Cl $3p$ -state in the crystal to a localized exciton state, which situation seems to roughly justify the approximate validity of the zero-order description of our exciton model.

For the numerical estimate, the explicit form of the Wannier function of a conduction band of KCl, i. e., $a_c(\mathbf{r} - \mathbf{R}_l)$ will be required but, unfortunately, we have no reliable knowledge of it at present. However, when we roughly assume a $4s$ atomic orbital of Slater type in Cl ion for $a_c(\mathbf{r} - \mathbf{R}_l)$, the exchange energy in (26) is found to become extremely small, whereas the Wannier function evaluated by a single OPW of a conduction band yields about $0.5 \sim 0.8$ eV for the mentioned exchange energy. Presumably,* its correct estimate will be supposed to lie near 0.1 eV or so. Furthermore, E^{exc} , estimated from our previous result on KCl and the half-empirical estimate of a conduction band bottom, becomes of $-0.5 \text{ eV} - 1.86 \text{ eV} = -2.36 \text{ eV}$. The values of $\varepsilon_n(\mathbf{R}_{ll'} - \mathbf{R}_l)$ estimated from the band width calculated by Howland⁽¹⁹⁾ and $S_{ll'}$ from the previous work on KCl lead to approximately 0.75 eV for $\sum S_{ll'} \varepsilon_n(\mathbf{R}_{ll'} - \mathbf{R}_l)$ and finally we get $\varepsilon_n(0) \simeq E_{\text{atom}}(3p) + V_M = -11.3 \text{ eV}$. It is quite difficult to estimate $\mathcal{J}E$ on a simple way as above although it is reasonably expected to yield a minor contribution. Substituting the above estimates of the various quantities in (26), we finally obtain

$$h\nu_0 = E^{(n,n)}(\mathbf{K}=0) - E_g \simeq 8.3 \text{ eV} + \mathcal{J}E \quad (28)$$

for a singlet exciton peak of KCl at $T=0^\circ\text{K}$.

According to the reflection measurement of Hartman, Nelson and Siegfried,⁽¹⁷⁾ the peak frequency of an exciton absorption in KCl has been observed to become 7.7 eV at

* This is the case because of the polarization effect of the crystal in addition to the improvement of the Wannier function.

$T=320^\circ\text{K}$ and then the peak frequency has been found to shift toward the shorter wave length side with decreasing temperatures. Therefore, the above estimate of peak frequency of KCl at $T=0^\circ\text{K}$ may be considered to be reasonable, allowing for the crude estimates of the various quantities and the contribution from JE . The more accurate evaluation of the exciton peak shall be reserved until the Wannier functions belonging to the conduction band can be obtained more accurately.

b) *Comparison of the present theory with that of Slater and Shockley⁵⁾.*

Finally, some remarks should be added to the comparison of the present results obtained above with those of Slater and Shockley. In the present method of procedure to treat an exciton, the Coulomb interaction of an excited electron with a positive hole around a halogen ion has been dealt with in the zero-order approximation of the formalism, whereas, in the theory of Slater and Shockley, the mentioned interaction has been worked in the next higher approximation. The above situation shows that the former method of treatment is more appropriate for a small size exciton and the latter one, for a large size exciton, respectively. In fact, the exciton wave function²³⁾ of type n and wave number vector \mathbf{K} may be written, in Slater-Shockley's theory, as

$$\Psi^{\mathbf{K},n} = \sum_{\mathbf{k}_e, \mathbf{k}_h, j, j'} \Psi_{\mathbf{k}_e, \mathbf{k}_h, j, j'}^{\mathbf{K},n} B(\mathbf{k}_e j; \mathbf{k}_h j') \quad (29)$$

in which $B(\mathbf{k}_e j; \mathbf{k}_h j')$ represents the determinantal function corresponding to a configuration with an electron with \mathbf{k}_e in a conduction band j and with a positive hole with \mathbf{k}_h in valence band j' . The states involved are restricted by the relation $\mathbf{k}_e + \mathbf{k}_h = \mathbf{K}$. The Fourier transform of $\Psi_{\mathbf{k}_e, \mathbf{k}_h, j, j'}^{\mathbf{K},n}$ may be written

$$(N\Omega)^{-1/2} e^{i\mathbf{K} \cdot \mathbf{R}} \phi_{jj'}^{\mathbf{K},n}(\mathbf{r}) \quad (30)$$

where \mathbf{R} is the radius vector of a center of mass of an electron-hole pair, \mathbf{r} the relative position vector of the pair, Ω the volume of a unit cell and N the number of such cells in the crystal. $\phi_{jj'}^{\mathbf{K},n}(\mathbf{r})$ expresses the internal state of an electron-hole pair and satisfies a Schrodinger-like-equation in the effective mass approximation. The first factor on the right-hand side of (29) represents the wave function of an exciton's over-all motion in the crystal. Superimposed upon this is a modulation produced by the atom cores of the crystal, whose effect is included in the Bloch orbitals involved in the second factor on the right-hand side of (29). The corresponding energies become, for a simple case in the effective mass approximation,

$$E = JE_F + \frac{\hbar^2 K^2}{2(m_e + m_h)} - \frac{\mu e^4}{2\hbar^2 \epsilon^2 n^2}, \quad (31)$$

where JE_F expresses the energy gap between a conduction and a valence band, m_e and m_h the effective masses of an electron in conduction band and of a hole in valence band, respectively. μ denotes the reduced mass, ϵ the dielectric constant and $n=1, 2, \dots$. The second term in (31) represents the translational energy of a center of mass of an exciton and the third the internal energy of an exciton. On the other hand, our wave function of an exciton in the present formalism may be written as

$${}^i\Psi = \sum_n {}^iA_n \sum_l e^{i\mathbf{K}\cdot\mathbf{R}_l} {}^i\Psi^{(nl,el)} \quad (32)$$

In accordance with the rather strong interaction of an excited electron with a hole, the internal motion of an exciton is not separated out from the influences of the atom cores in the crystal but is involved in the determinantal function ${}^i\Psi^{(nl,el)}$, although the translational part of an exciton is given by the factor, $\exp(i\mathbf{K}\cdot\mathbf{R}_l)$, in accordance with the situation that the translational motion of an exciton has been taken into account in the higher order approximation. The corresponding energy for a special case of a single valence band is given by (16). Furthermore, each term of (16) has fairly simple interpretation on the atomic view-point, as described before. Thus, we may conclude that the present method of procedure is reasonably valid for the case of a rather strong interaction of an excited electron with a positive hole and provides some refinement of the original work of Frenkel.

c) *Spin-orbit coupling.*

Now, we shall discuss qualitatively the effect of spin-orbit coupling on the structure of exciton multiplet. As a consequence of the approximations we have made, the spin-orbit coupling energy⁽²⁴⁾ of N valence electrons in the ion core field of the crystal may be written as

$$H_{SL} = -\frac{\hbar}{4m^2c^2} \sum_{j=1}^N [\text{grad} V_p(\mathbf{r}_j) \times \mathbf{P}_j] \cdot \boldsymbol{\sigma}_j + \frac{e^2\hbar}{2m^2c^2} \sum_{i \neq j} \frac{1}{r_{ij}^3} [\mathbf{r}_{ij} \times \mathbf{P}_j] \cdot \boldsymbol{\sigma}_i, \quad (33)$$

where $\mathbf{P} = \hbar/i \cdot \text{grad}$, $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ and the other quantities have the usual meaning.

Then, the Schroedinger wave equation becomes

$$(\mathbf{H} + H_{SL}) \cdot \Psi = E \cdot \Psi, \quad (34)$$

in which \mathbf{H} is given by (2). Our wave function is approximately represented by a linear combination of the triplet and singlet wave functions which satisfy (2).

$$\Psi = \sum_{i=1}^3 \{ {}^1a_i {}^1\Psi_i + {}^3a_{0i} {}^3\Psi_{0i} + {}^3a_{1i} {}^3\Psi_{1i} + {}^3a_{-1i} {}^3\Psi_{-1i} \}, \quad (35)$$

where

$$\left. \begin{aligned} {}^1\Psi_i &= \sum_n {}^1A_{ni} \sum_l e^{i\mathbf{K}\cdot\mathbf{R}_l} {}^1\Psi^{(nl,el)} \\ {}^3\Psi_{0i} &= \sum_n {}^3A_{0ni} \sum_l e^{i\mathbf{K}\cdot\mathbf{R}_l} {}^3\Psi_0^{(nl,el)} \\ {}^3\Psi_{1i} &= \sum_n {}^3A_{1ni} \sum_l e^{i\mathbf{K}\cdot\mathbf{R}_l} {}^3\Psi_1^{(nl,el)} \\ {}^3\Psi_{-1i} &= \sum_n {}^3A_{-1ni} \sum_l e^{i\mathbf{K}\cdot\mathbf{R}_l} {}^3\Psi_{-1}^{(nl,el)} \end{aligned} \right\}. \quad (36)$$

The substitution of (35) in (34) leads, in a straightforward way, to a 12×12 secular equation, which is so complicated to be soluble at present. Then the exciton multiplet is seen to consist of twelve components at most although their energy separations can hardly be predicted without the actual solution of the secular equation. The relative intensities of the exciton absorption components can be computed easily once the secular equation is solved and the corresponding coefficients 1a_i , ${}^3a_{0i}$, ${}^3a_{1i}$ and ${}^3a_{-1i}$ are

actually determined. In this case, since the singlet wave function is superimposed on the triplet one, the dipole transition to the mostly triplet exciton state from the ground state will be realized, resulting into the further splitting of the exciton peak. However, it should be mentioned that all the exciton multiplet components are not always observed experimentally on account of the limited resolution of optical instrument and the temperature-broadening. In order to see in more detail the gross behaviour of the effect of small spin-orbit coupling in alkali chlorides and fluorides, we shall actually compute the diagonal elements of the energy matrix involved in the secular equation or the first order energy-shift of the singlet and triplet exciton states corresponding to $\mathbf{K}=0$ (exciton peak) by the conventional perturbation method. The computations are straightforward and then the singlet exciton peak remains unchanged in energy, to this order of approximation. On the other hand, the triplet states ($\mathbf{K}=0$) corresponding to $M_s = \pm 1$ shift upward and downward by an equal amount, respectively, with respect to a triplet state $M_s=0$, which is unchanged to the same approximation, M_s being magnetic quantum number of the total spin. The corresponding coefficients of mixing singlet and triplet wave functions become proportional to the matrix elements of \mathbf{H}_{sl} with respect to singlet and triplet exciton states.

In their reflection method of measurement, Hartman, Nelson and Siegfried⁷ have actually observed the apparent splitting into doublet of the exciton absorption peak, at the temperatures below 150°K, in the cleaved crystals of KCl and NaCl. Their observed splitting at 6°K becomes of 0.14 eV in KCl, i. e., about a doublet separation of a chlorine atom (0.11 eV). The result would probably be interpreted as follows. In accordance with the relatively small spin-orbit coupling energy of the valence electrons in alkali chlorides, the singlet exciton peak will practically be unchanged even if we allow for the spin-orbit coupling and, furthermore, the splitting of the triplet exciton peak will perhaps be too small to be resolved experimentally in addition to the temperature broadening. The only sizable effect of the spin orbit coupling is to superimpose, in some measure, the singlet exciton wave function onto the triplet wave functions, which will presumably enable one of having non-vanishing but weak transition to the mostly triplet exciton state from the ground state (singlet state). If the above considerations were really valid, the shorter wave length peak of high intensity would be ascribed to the transition to a mostly singlet exciton state, while, the longer wave length peak of low intensity would be ascribed to the transition to a mostly triplet exciton state. According to the equation (26), therefore, the energy separation between two peaks will be approximately given by twice the exchange energy of an exciton-electron with a positive hole with the same location. Then, it follows

$$(el, nl|e^2/r_{12}|nl, el) \simeq 0.07 \text{ eV},$$

which, in turn, leads to $h\nu_0 \sim 8.3 \text{ eV}$ for a singlet exciton peak at $T=0^\circ\text{K}$.

Finally, the authors would like to express their gratitude for financial support from Grant-in-Aid of the Education Ministry of Japan.

References

- 1) H. Haken, Der heutige Stand der Exzitonenforschung in Halbleitern (Halbleiterprobleme IV).
- 2) J. Frenkel, Phys. Rev. **37** (1931), 17, 1276; ZS. d. Sowjetunion, **8** (1936), 185; **9** (1936), 158; J. Exp. Theor. Phys. **6** (1936), 647.
R. Peierls, Ann. Phys. **31** (1932), 905.
W. R. Heller and A. Markus, Phys. Rev. **84** (1951), 809.
R. S. Knox, Exciton States in Solid Argon, AFOSR. TN-58-370 (1958).
- 3) A. W. Overhauser, Phys. Rev. **101** (1956), 1702.
- 4) von Hippel, ZS. f. Phys. **101** (1936), 680. de Boer, *Electron Emission and Adsorption Phenomena* (Cambr. 1935), p. 242. N. F. Mott and R. W. Gurney; *Electronic Processes in Ionic Crystals*. (1940).
- 5) J. C. Slater and W. Shockley, Phys. Rev. **50** (1936), 705. J. C. Slater, Tech. Rep. No. 6 (1954), 3-31.
G. Wannier, Phys. Rev. **52** (1937), 191.
Y. Takeuchi, Prog. Theor. Phys. **18** (1957), 421.
- 6) J. Meyer, Physica, **22** (1956), 109.
- 7) J. M. Luttinger and W. Kohn, Phys. Rev. **97** (1955), 896, 883.
- 8) J. C. Slater and G. F. Koster, Phys. Rev. **95** (1954), 1167, 1436.
- 9) G. Dresselhuus, J. Phys. Chem. Solids, **1** (1956), 15.
- 10) F. Seitz, *The Modern Theory of Solids*, (1940).
- 11) T. Muto and H. Okuno, J. Phys. Soc. Japan, **11** (1956), 633, which is referred to Part I in the text. Report of Inst. Sci. Tech. Univ. Tokyo, **10** (1956), 17.
- 12) D. L. Dexter, Phys. Rev. **83** (1951), 435, **108** (1957), 707.
- 13) C. J. Delbecq, P. Pringsheim and P. Yuster, J. Chem. Phys. **19** (1951), 574; **20** (1952), 746.
W. Martienssen, ZS. f. Phys. **131** (1952), 488.
W. Martienssen and R. W. Pohl, ZS. f. Phys. **133** (1952), 153.
H. Ruchardt, ZS. f. Phys. **140** (1955), 547.
- 14) L. G. Parratt and E. L. Jossem, Physica, **XX** (1954), 1134; Phys. Rev. **97** (1955), 916.
- 15) D. Pines, *Solid State Physics*, Vol. 1, p. 376 (1955).
- 16) E. G. Schneider and H. M. O'Bryan, Phys. Rev. **51** (1937), 293.
R. Hilsch and R. W. Pohl, ZS. f. Phys. **59** (1930), 812.
- 17) J. R. Nelson, J. W. Siegfried and P. L. Hartman, Phys. Rev. **105** (1956), 123.
- 18) L. G. Parratt and E. L. Jossem, J. Phys. Chem. Solids, **2** (1957), 67.
- 19) L. P. Howland, Phys. Rev. **109** (1958), 1927. We should like to thank Dr. Howland for sending a detailed preprint of his results of calculation on KCl before publication.
H. Bethe, Handb. d. Phys. **24/2**, p. 404 (1933).
- 20) F. Bassani and N. Inchausti, Phys. Rev. **105** (1957), 819.
- 21) D. R. Hartree and W. Hartree, Proc. Roy. Soc. **A156** (1936), 45.
- 22) H. Bethe, Handb. d. Phys. **24/2**, p. 404 (1933).
- 23) G. Wannier, loc. cit. R. J. Elliot, Phys. Rev. **108** (1957), 1384.
- 24) H. Bethe, Handb. d. Phys. **24/1**, p. 377 (1933).

A Theory of Clothed Unstable Particles, II

Noboru NAKANISHI

Department of Physics, Kyoto University, Kyoto

(Received August 4, 1958)

The physical state of unstable particle previously proposed is derived as the state produced in a finite time interval on the basis of the scattering theory of Gell-Mann and Goldberger, and also the exact state is obtained from the physical state by employing an appropriate limiting process. An S -matrix-theoretical calculation method is proposed for the processes including unstable particles.

§ 1. Introduction

In the previous paper,¹⁾ which is quoted as I, we investigated the clothed state of unstable particles. By introducing a new mathematical notion, "complex distribution", we could construct the exact eigenstate of total Hamiltonian corresponding to the unstable particle, which we called "exact state of the unstable particle". But it being not observable, we defined the physical state of the unstable particle by its approximate state. Hence the definition of the physical state might seem rather artificial, though it was shown that this state satisfies various physical requirements. We did not clarify the physical meaning of the rather indefinite function $f(\omega)$ which appeared in the physical state.

Recently Naito²⁾ has investigated the finite-time scattering theory by the wave-packet method, and succeeded in treating the production and decay of the unstable particle. So, if we treat only the production in a finite time interval, it is expected that we can obtain the physical state of the unstable particle. In order to formulate this mathematically, we make use of Gell-Mann and Goldberger's scattering formalism.³⁾ We define the physical state by the state produced in a finite time interval $1/\kappa$, employing the more general form $\gamma_\kappa(T)$ instead of their switching function $\lim_{\epsilon \rightarrow +0} \epsilon e^{\epsilon T}$. We can then prove that it coincides with the physical state defined in I apart from the normalization constant, and that $f(\omega)$, which is now represented as a functional of $\gamma_\kappa(T)$, satisfies the four conditions postulated in I.

Further, we demonstrate that for $\kappa \rightarrow 0$ the above physical state tends to the exact state defined in I when $\gamma_\kappa(T)$ is chosen appropriately. The former thus is a good approximation of the latter for appropriate $\gamma_\kappa(T)$ with small κ .

It may be natural to use the physical state in the calculation of matrix elements relating to unstable particles. But it contains a rather indefinite function $f(\omega)$ which cannot

generally be controlled by experiments, and so the calculations become futilely complicated. We therefore propose to use the exact state instead of the physical one in the calculations including unstable particles. According to this method, we can treat transition matrices of the decay or scattering of unstable particles in the same way as in the usual S -matrix theory of stable particles. This method thus permits to treat unstable particles by the S -matrix-theoretical formalism.

§ 2. Physical state

In this section we employ Lee's model¹⁾ for the sake of simplicity. Notations are same as in I throughout this paper, unless otherwise indicated.

According to the scattering formalism of Gell-Mann and Goldberger, the outgoing-wave eigenstate is obtained by the following adiabatic process :

$$|N\theta(\mathbf{p})^+\rangle = \lim_{\varepsilon \rightarrow +0} \int_{-\infty}^0 dT \varepsilon e^{\varepsilon T} e^{i(H-m_N-\omega_p)T} \alpha_p^* |N\rangle. \quad (2.1)$$

This represents a superposition of the incident waves emitted in various times at the presence of interaction. Quite analogously to this, let us consider the following state :

$$\Psi^\kappa \equiv \int_{-\infty}^0 dT \gamma_\kappa(T) e^{i(H-E_0)T} |V\rangle, \quad (2.2)$$

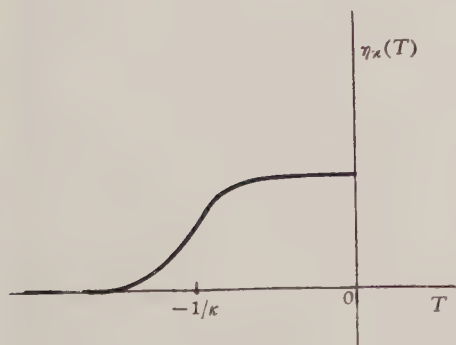


Fig. 1.

which represents the superposition of the V -particle amplitudes produced with intensity $\gamma_\kappa(T)$. Actually, the physical V -particle is produced in a finite time interval, and so the state Ψ^κ defined above may be regarded as the physical unstable state. Here we assume that the function $\gamma_\kappa(T)$ roughly takes such a form as Fig. 1. As κ must be finite, Ψ^κ generally depends on the form of the function $\gamma_\kappa(T)$. The usual choice

$$\gamma_\kappa(T) = \kappa e^{\kappa T} \quad (2.3)$$

is not preferable, because it has the following defects.

- i) It has a finite tangent at the final time $T=0$.*
- ii) It does not damp sufficiently rapidly for $T \rightarrow -\infty$.

The meaning of these respects will be clarified later. E_0 , which represents the rest energy of the V -particle, might be m_0 for the bare state, but we should use here the renormalized mass as E_0 , as is well known. Namely, we put

$$E_0 = m_V - i\gamma/2 \quad (2.4)$$

* This physically means that the production does not become stationary at the final time.

which is the pole of the modified propagator $S_V'(E)$. Hence $\gamma_\kappa(T)$ must damp faster than at least $e^{\tau T/2}$ in order to make the integral (2.2) converge. In order to calculate (2.2), we insert into it the complete set^{6,7)}

$$|N\theta(\mathbf{p})^+\rangle = g(\omega_p) S_V'(m_N + \omega_p) |V\rangle \\ + \iint \left[\delta(\mathbf{k} - \mathbf{p}) - g(\omega_p) S_V'(m_N + \omega_p) \frac{g(\omega_k)}{\omega_k - \omega_p - i\epsilon} \right] \alpha_k^* d\mathbf{k} |N\rangle \quad (2.5)$$

with

$$g(\omega_p) \equiv g \cdot G(\omega_p) \quad 1/2\omega_p.$$

We then obtain

$$\Psi^\kappa = \int_{-\infty}^0 \gamma_\kappa(T) dT \left[\int d\mathbf{p} \rho(\omega_p) e^{i(m_N + \omega_p - E_0)T} |V\rangle \right. \\ \left. + \int d\mathbf{k} g(\omega_k) S_V'^*(m_N + \omega_k) e^{i(m_N + \omega_k - E_0)T} \alpha_k^* d\mathbf{k} |N\rangle \right. \\ \left. - \int d\mathbf{p} \cdot \rho(\omega_p) e^{i(m_N + \omega_p - E_0)T} \int \frac{g(\omega_k)}{\omega_k - \omega_p - i\epsilon} \alpha_k^* d\mathbf{k} |N\rangle \right], \quad (2.6)$$

where

$$\rho(\omega_p) \equiv g^2(\omega_p) S_V'^*(m_N + \omega_p) S_V'(m_N + \omega_p). \quad (2.7)$$

By using Lehmann's spectral representation⁸⁾

$$S_V'^*(m_N + \omega_k) = \int d\mathbf{p} \frac{\rho(\omega_p)}{\omega_k - \omega_p - i\epsilon}, \quad (2.8)$$

(2.6) can be rewritten as

$$\Psi^\kappa = c|V\rangle + \int d\mathbf{p} \rho(\omega_p) \int F(\omega_k, \omega_p) g(\omega_k) \alpha_k^* d\mathbf{k} |N\rangle, \quad (2.9)$$

where

$$c \equiv \int d\mathbf{p} \rho(\omega_p) \int_{-\infty}^0 dT \gamma_\kappa(T) e^{i(m_N + \omega_p - E_0)T} \quad (2.10)$$

and

$$F(\omega_k, \omega_p) \equiv \int_{-\infty}^0 \gamma_\kappa(T) dT \frac{e^{i(m_N + \omega_k - E_0)T} - e^{i(m_N + \omega_p - E_0)T}}{\omega_k - \omega_p - i\epsilon}. \quad (2.11)$$

$F(\omega_k, \omega_p)$ is regular at $\omega_k = \omega_p$, and so $-i\epsilon$ in the denominator is unnecessary. $F(\omega_k, \omega_p)$ is rewritten as follows.

$$F(\omega_k, \omega_p) = -i \int_{-\infty}^0 \gamma_\kappa(T) e^{i(m_N + \omega_p - E_0)T} dT \int_T^0 e^{i(\omega_k - \omega_p)\tau} d\tau \\ = -i \int_{-\infty}^0 e^{i(m_N + \omega_k - E_0)\tau} d\tau \int_{-\infty}^\tau \gamma_\kappa(T) e^{i(m_N + \omega_p - E_0)(T - \tau)} dT$$

$$= -i \int_{-\infty}^0 dT e^{i(m_N + \omega_p - E_0)T} \int_{-\infty}^0 d\tau e^{i(m_N + \omega_k - E_0)\tau} \eta_{\kappa}(T + \tau). \quad (2.12)$$

Substituting (2.12) in (2.9) and integrating by parts relating to τ , we obtain

$$\begin{aligned} \mathcal{V}^{\kappa} &= c |V\rangle - \int d\mathbf{p} \rho(\omega_p) \int_{-\infty}^0 dT e^{i(m_N + \omega_p - E_0)T} \int d\mathbf{k} \frac{g(\omega_k)}{m_N + \omega_k - E_0} \\ &\quad \cdot \left[\eta_{\kappa}(T) - \int_{-\infty}^0 d\tau e^{i(m_N + \omega_k - E_0)\tau} \eta'_{\kappa}(T + \tau) \right] \alpha_k^* |N\rangle \\ &= c \left[|V\rangle - \int \frac{g(\omega_k)}{m_N + \omega_k - E_0} \{1 - f(\omega_k)\} \alpha_k^* d\mathbf{k} |N\rangle \right], \end{aligned} \quad (2.13)$$

where

$$f(\omega_k) \equiv \int_0^{+\infty} e^{-i(m_N + \omega_k - m_V)t} h(t) dt \quad (2.14)$$

with

$$h(t) \equiv c^{-1} e^{\gamma t/2} \int d\mathbf{p} \rho(\omega_p) \int_{-\infty}^0 dT \eta'_{\kappa}(T - t) e^{i(m_N + \omega_p - E_0)T}. \quad (2.15)$$

Here the function $h(t)$ has the following properties:

$$\begin{aligned} \text{i)} \quad \int_0^{+\infty} h(t) dt &= c^{-1} \int d\mathbf{p} \rho(\omega_p) \int_{-\infty}^0 dT \left[\eta_{\kappa}(T) + \int_0^{+\infty} (\gamma/2) e^{\gamma t/2} \eta_{\kappa}(T - t) dt \right] e^{i(m_N + \omega_p - E_0)T} \\ &= 1 + O(\gamma/\kappa), \end{aligned} \quad (2.16)$$

$$\text{ii)} \quad h(0) = c^{-1} \int d\mathbf{p} \rho(\omega_p) \int_{-\infty}^0 dT \eta'_{\kappa}(T) e^{i(m_N + \omega_p - E_0)T} \approx 0. \quad (2.17)$$

Because $\eta'_{\kappa}(T)$ is a sufficiently smooth function and has the values appreciably different from zero only in the neighbourhood of $T = -1/\kappa$, and so the integral approximately vanishes due to the rapidly oscillating factor. For $h'(0)$, $h''(0)$, etc., the same is true, provided that $\eta_{\kappa}(T)$ has the higher derivatives.

Now the final expression of \mathcal{V}^{κ} , (2.13), has completely the same form with the physical state $|V^f\rangle$ given in I, apart from the normalization factor. And, further, we can prove that $f(\omega_k)$ actually satisfies the four conditions postulated in I.

$$\text{i)} \quad f(\omega_k) \approx 1 \quad \text{for} \quad |\omega_k + m_N - m_V| \ll \kappa.$$

Because we then get $f(\omega_k) \approx \int_0^{+\infty} h(t) dt \approx 1$ from (2.16) for $\gamma \ll \kappa$.

$$\text{ii)} \quad f(\omega_k) \approx 0 \quad \text{for} \quad |\omega_k + m_N - m_V| \gg \kappa \quad \text{where} \quad \omega_k \text{ is real.}$$

Because, $f(\omega_k)$ then vanishes approximately owing to (2.17) for the contribution from $t \approx 0$ and to the rapidly oscillating factor for that from the other part.

$$\text{iii)} \quad f(\omega_k) \text{ is regular near the real axis.}$$

The analyticity in the lower half-plane is obvious from (2.14). Further, the

Fourier transform $b(t)$ nearly vanishes in the neighbourhood of $t=0$, and thus $f(\omega_k)$ is regular near the real axis.

iv) $f(\omega_k)$ vanishes far away in the lower half-plane.

This is obvious from (2.14).

Thus the physical meaning of the parameter κ in $|V\rangle$ is the inverse of the production time, as was expected.

Finally, it is of interest to consider the case of the simple but unpreferable choice $\eta_\kappa(T) = \kappa e^{\kappa T}$. In this case, the integration (2.11) is easily carried out, and the result is

$$\Psi^* = c \left[|V\rangle - \int \frac{g(\omega_k)}{m_N + \omega_k - E_0 - i\kappa} \alpha_k^* dk |N\rangle \right], \quad (2.18)$$

which is nothing but the state proposed by Glaser and Källén.* Then $f(\omega_k)$ becomes

$$f(\omega_k) = -i\kappa / (m_N + \omega_k - E_0 - i\kappa), \quad (2.19)$$

which decreases only linearly at a great distance and has a pole $\omega_k = E_0 - m_N + i\kappa$ near the real axis because of $\eta_\kappa'(0) = \text{finite}$.

§ 3. General case

In this section we treat the general case. As in I, we write $H = H_0 + H_1$, and denote the eigenstates of H_0 by $|V\rangle$ and $|n\rangle$. Then the eigenstate of H , $|n^+\rangle$, is given by*

$$|n^+\rangle = -i\varepsilon (H - E_n - i\varepsilon)^{-1} |n\rangle. \quad (3.1)$$

It should be noticed that the state (3.1) is *not* normalized when incident particles have their clouds.⁷⁾ We, nevertheless, treat it as if it were normalized, since this matter is not so important. The correct renormalization will be discussed in Appendix. Rewriting $(H - E_n - i\varepsilon)^{-1}$ in terms of U -matrix according to (4.5) in Araki *et al.*,⁸⁾ we obtain

$$\begin{aligned} |n^+\rangle &= \langle V | U(E_n) | n \rangle S_V'(E_n) |V\rangle + |n\rangle - \int dm \frac{\langle m | U(E_n) | n \rangle}{E_m - E_n - i\varepsilon} |m\rangle \\ &= \langle V | U(E_n) | n \rangle S_V'(E_n) \int dm \frac{\langle m | U(E_n) | V \rangle}{E_m - E_n - i\varepsilon} |m\rangle. \end{aligned} \quad (3.2)$$

Here $\langle V | U(E_n) | n \rangle$ corresponds to $g(\omega_n)$ in Lee's model, in which the third term of the right-hand side vanishes owing to the special type of interaction. Hence the derivation in last section does not hold as it is. We therefore newly put $H = \tilde{H}_0 + H_d$, where H_d is the decay interaction part of V -particle and \tilde{H}_0 is the other part. Using \tilde{H}_0 instead of H_0 , we have a similar expression as (3.2), but then the third term is removed. We

* In their paper, the sign of κ ($\equiv \lambda$) is indefinite, but here it is limited to $\kappa > 0$.

denote the corresponding quantities by affixing the symbol \sim . The calculations in last section can be straightforwardly repeated. But it must be noticed that the numerator of the integrand of the fourth term, $\langle \tilde{m} | \tilde{U}(E_n) | \tilde{V} \rangle$, depends not only on m but also on E_n . Namely, this E_n -dependence becomes a drag on factorizing c in (2.13). Hence, if we assume that $\langle \tilde{m} | \tilde{U}(E_n) | \tilde{V} \rangle$ does not strongly depend on E_n and so the main contribution comes out from the neighbourhood of $E_n = E_0$, then we can approximately replace it by $\langle \tilde{m} | \tilde{U}(E_0) | \tilde{V} \rangle$ and finally obtain

$$\Psi^\kappa \simeq c \left[|\tilde{V}\rangle - \int dm \frac{\langle \tilde{m} | \tilde{U}(E_0) | \tilde{V} \rangle}{E_m - E_0} \{1 - f(E_m)\} |\tilde{m}\rangle \right]. \quad (3.3)$$

Now, \tilde{H}_0 is not necessarily equal to $H_0 + H_s$. For example, for the weak interaction of $\Lambda N \pi$ the part $\Lambda \rightarrow N + \pi$ belongs to H_u but the part $\Lambda + \pi \rightarrow N$ belongs to \tilde{H}_0 . If such an unnatural division of H is not desirable, we should put $\tilde{H}_0 = H_0 + H_s$ and neglect the contribution of the third term caused by $H_w - H_d$, which is indeed very small.

These approximations employed above become better for smaller κ , and also depend on the form of the function $\gamma_\kappa(T)$. For instance, if $\gamma_\kappa(T) = \kappa e^{\kappa T}$, the approximations are not so good even for small κ , and it is not allowed to make κ smaller than $\gamma/2$. But if we adopt a very rapidly damping function like

$$\gamma_\kappa(T) = (2\kappa/\sqrt{\pi}) e^{-\kappa^2 T^2}, \quad (3.4)$$

then the approximations speedily become good. Further, in this case, we can define the limit $\kappa \rightarrow 0$, and Ψ^κ tends to the exact state defined in I, as is expected. We prove this in the following.

As a preliminary, we calculate the limit of an integral,

$$I \equiv \lim_{\kappa \rightarrow +0} \int_{-\infty}^0 \gamma_\kappa(T) e^{iET} dT, \quad (3.5)$$

where E is an arbitrary complex number, and $\gamma_\kappa(T)$ is given by (3.4). Carrying out the integration, we easily obtain the following result.

- i) $I = 0$ for $\text{Im } E < 0$.
- ii) $I = 0$ for $\text{Im } E \geq 0$ and $|\text{Re } E| > |\text{Im } E|$.
- iii) $I = 1$ for $E = 0$.
- iv) otherwise $I = \text{indefinite (or infinite)}$.

Next, making use of this, we evaluate a functional

$$I[\varphi] \equiv \lim_{\kappa \rightarrow +0} \int_{-\infty}^0 \gamma_\kappa(T) dT \int_M^{+\infty} e^{i(E-E_0)T} \varphi(E) dE. \quad (3.6)$$

Here $E_0 = m_F - i\gamma/2$ and $m_F - M > \gamma/2 + \delta$, $\delta > 0$ (infinitesimal). M physically represents

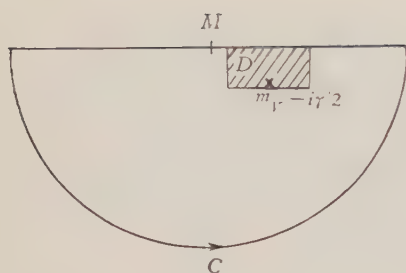


Fig. 2.

the threshold energy ($M = m_N + \mu$ in Lee's model). $\varphi(E)$ is a meromorphic function in the lower half-plane and regular in the domain D except for $E = E_0$ at which it may have a pole of the first order. The domain D is defined by

$$\delta > \text{Im } E > -(\gamma/2 + \delta). \quad (3.7)$$

Denoting the distant, large semi-circle in the

lower half-plane by C , we can write

$$I[\varphi] = \lim_{\kappa \rightarrow -0} \int_{-\infty}^0 \gamma_{\kappa}(T) dT \left[\left(\int_C + \int_M \right) e^{i(E-E_0)T} \varphi(E) dE - 2\pi i \sum_j e^{i(E_j-E_0)T} R_j \right], \quad (3.8)$$

where R_j stands for the residue of $\varphi(E)$ at E_j . The contributions from other than $E = E_0$ all vanish on account of i) and ii) presented above. We thus obtain

$$I[\varphi] = -2\pi i \lim_{E \rightarrow E_0} (E - E_0) \varphi(E). \quad (3.9)$$

Now, we calculate the limiting state, which stands for the idealized physical state produced in the infinite time interval,

$$\begin{aligned} \Psi &\equiv \lim_{\kappa \rightarrow -0} \Psi^* \\ &= \lim_{\kappa \rightarrow -0} \int_{-\infty}^0 \gamma_{\kappa}(T) dT \int dn e^{i(E_n - E_0)T} \langle n | U^*(E_n) | V \rangle S_V'^*(E_n) | n^+ \rangle, \end{aligned} \quad (3.10)$$

where $|n^+\rangle$ is given by (3.2). We analytically continue the integrand to the lower half-plane relating to E_n . We then know that $S_V'(E_n)$ has a pole of the first order at $E_n = E_0$, and assume that $S_V'(E_n)$ and U -matrix elements have no other singularities in the domain D . Then (3.9) assures that the second and third terms of (3.10) where $|n^+\rangle$ is expressed by (3.2) vanish, and we obtain

$$\Psi = c' \left[|V\rangle - \int dm \frac{\langle m | U(E_0) | V \rangle}{E_m^{(\kappa)} - E_0} |m\rangle \right], \quad (3.11)$$

using the notation of complex distribution.* Here

$$c' = -2\pi i Z_s^{-1} S_V'^*(E_0) \int dn \delta(E_n - E_0) \langle V | U(E_n) | n^+ \rangle \langle n | U^*(E_n) | V \rangle \quad (3.12)$$

with

$$Z_s^{-1} \equiv (\partial/\partial E) [S_V'(E)]^{-1} \Big|_{E=E_0} = \left[\lim_{E \rightarrow E_0} (E - E_0) S_V'(E) \right]^{-1}. \quad (3.13)$$

* Here the field operators are, of course, generalized to the complex distributional sense as in I. But, rather, we should understand this procedure as the analytic continuation (relating to E_n) of $\langle \psi(E_m) | m n^+ \rangle dE_m$, where $\psi(E_m)$ is an arbitrary regular function.

Using

$$[\mathcal{S}_V'^*(E_0)]^{-1} = \Sigma(E_0) - \Sigma^*(E_0) = \langle V | [U(E_0) - U^*(E_0)] | V \rangle \quad (3.14)$$

and

$$\begin{aligned} U(E) - U^*(E) &= H_1 [1 + (1 - A) (H_0^{(-)} - E)^{-1} H_1]^{-1} [(H_0^{(+)} - E)^{-1} - (H_0^{(-)} - E)^{-1}] \\ &\quad \cdot (1 - A) H_1 [1 + (1 - A) (H_0^{(-)} - E)^{-1} H_1]^{-1} \\ &= -2\pi i U^*(E) \delta(H_0 - E) (1 - A) U(E), \end{aligned} \quad (3.15)$$

we find

$$c' = Z_s. \quad (3.16)$$

(3.11) proves that \mathcal{V} exactly coincides with the exact state $|V\rangle$ defined in I apart from the coefficient Z_s . Though this result is the expected one, it is important that the relation between the time integral and the complex distribution of energy has been clarified. Namely, roughly speaking, the following conditions are necessary to yield the expected complex distribution.

- i) The switching function $\gamma_n(T)$ must damp very rapidly at the remote past, that is to say, the particle must be created more quickly than its decay.
- ii) The threshold energy must be sufficiently distant from the level of the unstable particle in comparison with the level width ($m_V - M \gg \gamma/2$).
- iii) There must be no other unstable levels near $m_V - i\gamma/2$ on the complex plane.

When these conditions are satisfied, the approximations employed in the derivation of the physical state are well justified. Indeed, for $\gamma_n(T)$ presented by (3.4), for example, the integration on T yields a factor $\exp[-\{(E_n - m_V)^2 - (\gamma/2)^2\}/4\kappa^2]$ approximately, and so the contribution from $|E_n - m_V| \gg \kappa$ becomes very small. Thus it suffices to consider only the contribution from the neighbourhood of $E_n = E_0$ in the integration on E_n .

§ 4. S-matrix theory of unstable particles

In the conventional theory the decay amplitude of the V -particle is calculated by

$$\langle n^- | V \rangle, \quad (4.1)$$

where $|n^- \rangle$ is the incoming-wave eigenstate of H , while the S -matrix is defined by

$$S_{nm} = \langle n^- | m^+ \rangle. \quad (4.2)$$

The asymmetry between the above treatments is obvious. The former is the transition amplitude from $t=0$ to $t=+\infty$, while the latter is that from $t=-\infty$ to $t=+\infty$. If we force ourselves to use the S -matrix for the decay in the framework of the conventional theory, we have

$$S_{nV} = -2\pi i \cdot \delta(E_n - m_V) \langle n^- | H_1 | V \rangle, \quad (4.3)$$

which exactly vanishes,⁹⁾ though the perturbation calculation yields divergent results (so-called "infrared-catastrophe-like divergence"¹⁰⁾) except for the lowest order.

Now, we have proposed in I that the decay amplitude should be calculated by

$$\langle n^- | V^f \rangle. \quad (4.4)$$

But the calculations using the physical state are generally very cumbersome. And since we cannot accurately control the production rate $\gamma_{\pi}(T)$ experimentally, detailed calculations are meaningless. Hence we propose to calculate the decay amplitude by

$$N \langle n^- | V \rangle, \quad (4.5)$$

where N is a normalization constant which is determined later, instead of (4.4). The approximation used here is very good as stated in § 3, though, of course, such a replacement as I (4.1) should be made for complex distribution in the evaluation of probability (cf. I (3.11) and I (4.9)). As was proved in § 3, (4.5) can be interpreted as the transition amplitude from $t = -\infty$ to $t = +\infty$. It was shown in I that $|V\rangle$ is apparently orthogonal to $|m^+\rangle$. When we introduce the "conjugate state of $|V\rangle$ ":

$$|V^*\rangle \equiv |V\rangle - \int dn \frac{\langle n | U^*(E_0^*) | V \rangle}{E_n^{(+)} - E_0^*} |n\rangle \quad (4.6)$$

with

$$H|V^*\rangle = E_0^*|V^*\rangle,$$

which is apparently orthogonal to $|n^-\rangle$, we have the "extended S -matrix":

$$\begin{pmatrix} \langle n^- | m^+ \rangle & N \langle n^- | V \rangle \\ N \langle V^* | m^+ \rangle & N^2 \langle V^* | V \rangle \end{pmatrix}. \quad (4.7)$$

We can easily prove

$$\langle V^* | V \rangle = Z_s^{-1}, \quad (4.8)$$

where Z_s is defined by (3.13), by means of the same technics as I (5.12) ~ (5.13). If we require that the last element of (4.7) is equal to unity, we obtain

$$N = Z_s^{1/2} \quad (4.9)$$

as the normalization constant of $|V\rangle$, accordingly, that of $|V^*\rangle$ is $N^* = Z_s^{1/2}$.

The (nV) -element of (4.7), $Z_s^{1/2} \langle n^- | V \rangle$, corresponds to the decay amplitude. We can now rewrite this in the R -matrix form³⁾:

$$\begin{aligned} Z_s^{1/2} \langle n^- | V \rangle &= -2\pi i \cdot \delta(E_n - E_0) Z_s^{1/2} \langle n | H_1 | V \rangle \\ &= -2\pi i \cdot \delta(E_n - E_0) Z_s^{1/2} \langle n | H_1 [1 - (1 - A)(H_0^{(-)} - E_0)^{-1} U(E_0)] | V \rangle \\ &= -2\pi i \cdot \delta(E_n - E_0) Z_s^{1/2} \langle n | U(E_0) | V \rangle. \end{aligned} \quad (4.10)$$

We thus see that the decay probability to $|n\rangle$ is given by

$$\frac{|f(E_n)|^2}{(E_n - m_\pi)^2 + (\gamma/2)^2} |Z_s^{1/2} \langle n | U(m_\pi - i\gamma/2) | V \rangle|^2. \quad (4.11)$$

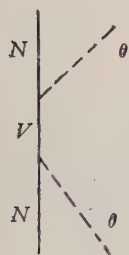


Fig. 3.

The total decay probability should be approximately normalized to unity, because this approximately represents the decay probability of the physical state. Hence we have

$$\gamma/2 \cong \pi \int dn \cdot \delta(E_n - m_V) |Z_s^{1/2} \langle n | U(m_V) | V \rangle|^2 \quad (4.12)$$

for very small γ . And we find that the transition probability per unit time for $\gamma \rightarrow 0$ is

$$2\pi \delta(E_n - m_V) |Z_s^{1/2} \langle n | U(m_V) | V \rangle|^2, \quad (4.13)$$

which is a well-known formula¹¹⁾ apart from the Z -factor.

Finally, we briefly discuss the scattering problems including unstable particles. The perturbational calculation of these processes usually lead to the so-called infrared-catastrophe-like divergence in the conventional theory. For example, consider the N - θ resonance scattering in Lee's model. Though the exact calculation of $\langle N\theta(\mathbf{p}')^- | N\theta(\mathbf{p})^+ \rangle$ does not give rise to any divergence, the free propagator of the V -particle is

$$S_V(E) = 1/(E - m_V + i\epsilon) \quad (4.14)$$

and the integral of its absolute square is divergent when the V -particle is unstable. Hence we should use

$$S_V(E) = 1/(E - E_0) \quad (4.15)$$

as the free propagator instead of (4.14).

The scattering of unstable particles can hardly be treated by the conventional theory. Then we apply to this problem the method of the extended S -matrix proposed above. For example, the S -matrix for the V - θ scattering in Lee's model is defined by

$$\langle V^*\theta(\mathbf{p}')^- | V\theta(\mathbf{p})^+ \rangle, \quad (4.16)$$

where

$$H | V\theta(\mathbf{p})^+ \rangle = (E_0 + \omega_p) | V\theta(\mathbf{p})^+ \rangle$$

and

$$H | V^*\theta(\mathbf{p})^- \rangle = (E_0^* + \omega_p) | V^*\theta(\mathbf{p})^- \rangle.$$

In the perturbational calculation the propagator of the N -particle in Fig. 4 is

$$S_N(m_V - \omega_p) = 1/(m_V - \omega_p - m_N + i\epsilon), \quad (4.17)$$

which leads to a divergence, while in the present method it becomes

$$S_N(E_0 - \omega_p) = 1/(E_0 - \omega_p^{(-)} - m_N), \quad (4.18)$$

and finally is replaced by

$$\{1 - f(\omega_p)\} / (E_0 - \omega_p - m_N). \quad (4.19)$$



Fig. 4.

Thus we can treat the scattering problems of unstable particles.

Acknowledgements

The author would like to express his sincere thanks to Messrs. K. Naito and T. Kanki and Prof. Y. Munakata for valuable discussions. He is also indebted to the financial aids of a Yukawa Fellowship of Osaka University and a Kumura Scholarship of Teikoku Rayon Co., Ltd.

Appendix

The state (3.1) is not normalized, *i.e.*

$$\langle n^+ | n^+ \rangle = Z_n,$$

where Z_n is the product of the Z -factors of incident particles. Accordingly, the completeness should be revised as

$$\int dn |n^+\rangle Z_n^{-1} \langle n^+| = 1.$$

Hence the integrand of (3.12) contains an additional factor Z_n^{-1} . The calculation (3.15) which is usually accepted⁷ is no longer valid. To understand this, consider, for example, its fourth order terms of perturbation:

$$\begin{aligned} & H_1 [(1-A)(H_0^{(+)}-E)^{-1}H_1]^3 - H_1 [(1-A)(H_0^{(-)}-E)^{-1}H_1]^3 \\ &= H_1 (1-A) [(H_0^{(+)}-E)^{-1} - (H_0^{(-)}-E)^{-1}] H_1 [(1-A)(H_0^{(+)}-E)^{-1}H_1]^2 \\ &+ H_1 (1-A)(H_0^{(-)}-E)^{-1}H_1 (1-A) [(H_0^{(+)}-E)^{-1} - (H_0^{(-)}-E)^{-1}] \\ &H_1 (1-A)(H_0^{(+)}-E)^{-1}H_1 \\ &+ H_1 [(1-A)(H_0^{(-)}-E)^{-1}H_1]^2 (1-A) [(H_0^{(+)}-E)^{-1} - (H_0^{(-)}-E)^{-1}] H_1. \end{aligned}$$

If $(1-A)H_1(1-A)(H_0^{(\pm)}-E)^{-1}H_1(1-A)$ has diagonal elements (*i.e.* self-energies), then the products $(H_0^{(+)}-E)^{-1} \dots (H_0^{(-)}-E)^{-1}$ appearing in the intermediary terms of the right-hand side are not well defined. Thus (3.15) is inadequate.

We will investigate this problem by means of the following concrete example ("doubled Lee model") which is exactly soluble. The Hamiltonian is given by

$$\begin{aligned} H = & m_V^0 \zeta_V^* \zeta_V + m_N^0 \zeta_N^* \zeta_N + m_M \zeta_M^* \zeta_M + \int \omega_k \theta_k^* \theta_k dk + \int \omega_k \phi_k^* \phi_k dk \\ & + \int g(\omega_k) dk [\zeta_V^* \zeta_N \theta_k + \zeta_V \zeta_N^* \theta_k^*] + \int f(\omega_k) dk [\zeta_N^* \zeta_M \phi_k + \zeta_N \zeta_M^* \phi_k^*]. \end{aligned}$$

Not only the V -particle but also the N -particle have their clouds. Their modified propagators respectively are

$$[S_V'(E)]^{-1} = E - m_V^0 + i\varepsilon - \int g^2(\omega_k) S_N'(E - \omega_k) dk,$$

$$[S_N'(E)]^{-1} = E - m_N^0 + i\varepsilon - \Sigma_N(E),$$

where

$$\Sigma_N(E) \equiv \int \frac{f^2(\omega_k) d\mathbf{k}}{E - m_N - \omega_k + i\varepsilon}.$$

The elements of the U -matrix are as follows.

$$\begin{aligned} \langle V|U(E)|V\rangle &= \int g^2(\omega_k) S_N'(E - \omega_k) d\mathbf{k}, \\ \langle V|U(E)\theta_p^*|N\rangle &= g(\omega_p) \cdot (E - m_N - \omega_p + i\varepsilon) S_N'(E - \omega_p), \\ \langle N|\theta_{p'} U(E)\theta_p^*|N\rangle &= \\ &= (\Sigma_N(E - \omega_p) - \delta m_N) (E - m_N - \omega_p + i\varepsilon) S_N'(E - \omega_p) \delta(\mathbf{p} - \mathbf{p}'), \\ \langle M|\theta_{p'} \phi_k U(E)\theta_p^*|N\rangle &= f(\omega_k) \cdot (E - m_N - \omega_p + i\varepsilon) S_N'(E - \omega_p) \delta(\mathbf{p} - \mathbf{p}'), \\ \langle V|U(E)\theta_p^* \phi_k^*|M\rangle &= g(\omega_p) f(\omega_k) S_N'(E - \omega_p). \end{aligned}$$

Hence we obtain

$$\begin{aligned} (1 - A) (m_N + \omega_p - H_0 + i\varepsilon)^{-1} U (m_N + \omega_p) \theta_p^* |N\rangle \\ = Z_N (1 - Z_N^{-1}) \theta_p^* |N\rangle + Z_N \int \frac{f(\omega_k)}{m_N - m_M - \omega_k + i\varepsilon} \theta_p^* \phi_k^* d\mathbf{k} |M\rangle, \\ (1 - A) (m_N + \omega_p - H_0 + i\varepsilon)^{-1} U (m_N + \omega_p) A S_V' (m_N + \omega_p) U (m_N + \omega_p) \theta_p^* |N\rangle \\ = Z_N g(\omega_p) S_V' (m_N + \omega_p) \left[\int g(\omega_k) S_N' (m_N + \omega_p - \omega_k) \theta_k^* d\mathbf{k} |N\rangle \right. \\ \left. + \iint \frac{f(\omega_k) f(\omega_{k'}) S_N' (m_N + \omega_p - \omega_k)}{m_N + \omega_p - m_M - \omega_k - \omega_{k'} + i\varepsilon} \theta_k^* d\mathbf{k} \phi_{k'}^* d\mathbf{k}' |M\rangle \right], \end{aligned}$$

where

$$Z_N \equiv \lim_{E \rightarrow m_N} (E - m_N) S_N'(E),$$

and thus from (3.2)

$$\begin{aligned} Z_N^{-1} |N\theta(\mathbf{p})^+\rangle &= g(\omega_p) S_V' (m_N + \omega_p) |V\rangle \\ &+ \int [\delta(\mathbf{k} - \mathbf{p}) + g(\omega_p) S_V' (m_N + \omega_p) g(\omega_k) S_N' (m_N + \omega_p - \omega_k)] \theta_k^* d\mathbf{k} |N\rangle \\ &+ \iint \frac{f(\omega_k') [\delta(\mathbf{k} - \mathbf{p}) + g(\omega_p) S_V' (m_N + \omega_p) g(\omega_k) S_N' (m_N + \omega_p - \omega_k)]}{m_N + \omega_p - m_M - \omega_k - \omega_{k'} + i\varepsilon} \\ &\quad \times \theta_k^* d\mathbf{k} \phi_{k'}^* d\mathbf{k}' |M\rangle \end{aligned}$$

and

$$\langle N\theta(\mathbf{p}')^+ | N\theta(\mathbf{p})^+ \rangle = Z_N \cdot \delta(\mathbf{p} - \mathbf{p}').$$

Now, we compare the both sides of (3.15).

$$\begin{aligned}
\langle V|[U(E) - U^*(E)]|V\rangle &= 2i \int g^2(\omega_k) \text{Im} S_N'(E - \omega_k) dk \\
&= -2\pi i \int g^2(\omega_k) [Z_N \cdot \delta(E - m_N - \omega_k) + \sigma(E - \omega_k)] dk, \\
-2\pi i \langle V|U^*(E) \delta(H_0 - E) (1 - A) U(E)|V\rangle \\
&= -2\pi i \int g^2(\omega_k) [Z_N^2 \cdot \delta(E - m_N - \omega_k) + \sigma(E - \omega_k)] dk,
\end{aligned}$$

where

$$\sigma(E) \equiv |S_N'(E)|^2 \int f^2(\omega_{k'}) \delta(E - m_N - \omega_{k'}) dk'.$$

Thus the first terms of the both do not coincide. The two corrections in (3.12) and in (3.15) are cancelled each other, and hence the final result (3.16) is unchanged, as it should be.

References

- 1) N. Nakanishi, *Prog. Theor. Phys.* **19** (1958), 607.
- 2) K. Yamamoto, **20** (1958), 857.
- 3) M. Gell-Mann and M. L. Goldberger, *Phys. Rev.* **91** (1953), 398.
B. A. Lippmann and J. Schwinger, *Phys. Rev.* **79** (1950), 469.
- 4) T. D. Lee, *Phys. Rev.* **95** (1954), 1329.
- 5) H. Lehmann, *Nuovo Cim.* **11** (1954), 342.
See also references 1) and 8).
- 6) V. Glaser and G. Källén, *Nuclear Phys.* **2** (1957), 706.
- 7) B. S. De Witt, *Phys. Rev.* **100** (1955), 905.
T. Kanki, *Nuovo Cim.* **6** (1957), 628.
- 8) H. Araki, Y. Munakata, M. Kawaguchi and T. Goto, *Prog. Theor. Phys.* **17** (1957), 419.
- 9) K. Naito, *Prog. Theor. Phys.* **18** (1957), 200.
N. Nakanishi, *Soryushiron-Kenkyu* (mimeographed circular in Japanese) **16** (1957), 164.
- 10) D. Ito, *Prog. Theor. Phys.* **6** (1951), 1020 (L).
N. Nakanishi, *Soryushiron-Kenkyu* (mimeographed circular in Japanese) **15** (1957), 344.
- 11) W. Heitler, *The Quantum Theory of Radiation*, 3rd ed. (1954, Oxford).

Note added in proof 1) Prof. K. Kunugui and others kindly informed to the author that a distribution of analytic function was defined by Köthe [*Math. Zeit.* **57** (1952), 13], though it rather differs from our complex distribution.

2) Recently T. Goto also has derived an exact state of unstable particle in a quite different way. This state is identical with ours, but his interpretation does not coincide with ours.

The Dispersion Relation and the Pion-Nucleon Interaction Cross Section at Very High Energies

Tetsuro SAKUMA, Shinya FURUI and Akira KANAZAWA

Department of Physics, Hokkaido University, Sapporo

(Received July 28, 1958)

Using the dispersion relations, the sum of the total cross sections for the π^+p and π^-p interactions are evaluated at very high energies, under the assumption that the cross sections are constants above a certain very high energy. The numerical value of the sum of the cross sections is determined mainly by the empirical data at low energies (less than 2Bev). The analysis gives the sum of the cross sections a few times as large as twice of the geometrical cross section obtained by assuming that the "radius" of the nucleon is the pion Compton wavelength. Since the dispersion relations seem to be consistent with the experiments at low energies, this strongly suggests that the dispersion relation would break down at very high energies.

§ 1. Introduction

It is very interesting to investigate the consistency between the dispersion relations for the pion-nucleon scattering amplitudes and the experiments, because the inconsistency between them means the break down of at least one of the fundamental assumptions which are necessary to write down the dispersion relations, for example, the micro-causality, the charge-symmetry, the conservation of the probability and so on. At low energies this problem has been studied by many authors.¹⁾ The experiments seem to favour the dispersion relation, though the conclusions are not yet decisive. The reason is that the dispersion relation involves a principal integral and then the small experimental errors yield large ambiguities²⁾ as long as we restrict our attention only in low energy regions. Thus it becomes very interesting to analyse the dispersion relation at high energies. Furthermore, it is very likely that *the discrepancy would appear at high energies if the dispersion relation were inconsistent with the experiments.*

In this paper we investigate the total cross section at very high energies ($\gtrsim 30$ Bev), using the new dispersion relation given by W. Gilbert,³⁾ under the following assumptions;

I) The sum of the total cross sections becomes a constant at a certain high energy, say Z_H .

II) It changes linearly as the energy varies from 2 Bev to Z_H . The second assumption is introduced just for facilitating the calculation, so not an essential one. The integral to be calculated is fairly convergent and the contributions to the integral from this energy range are very small. The first assumption is plausible for two reasons; first, the energy dependence of the total cross section is theoretically limited by the relation.⁴⁾

$$\sigma(\omega) \propto \omega^n, \quad -1 \leq n < 1,$$

at high energies. Secondly, empirically it seems to be a constant (~ 30 mb) above the energy about 2 Bev.⁵⁾

Following Gilbert, we can express the sum of the total cross sections for the $\pi^- - p$ and $\pi^- - p$ reactions at a certain high energy, say Z_K , as the sum of the coupling constant term, the cross sections at the threshold and the convergent integral involving the real parts of the scattering amplitudes.⁶⁾ Under the assumptions the integral can be evaluated if we use the ordinary dispersion relation to obtain the real parts in the energy region where the empirical phase shifts are not available (above 300 Mev). *This partial use of the ordinary dispersion relation is the essential point of the present paper.* Here it should be noted that we must not make Z_K infinite before performing the integration, otherwise we would obtain only the vanishing cross sections at the infinite energy. The reason is that we have only the identity of the cross sections at the energy Z_K , if we substitute the integral of the imaginary parts for the real parts in all the energy regions. Then, to make Z_K infinite before the integration means to discard the one side of the identity.*

The sum of the total cross sections thus obtained is fairly insensitive to the choice of Z_K but depends on the energy Z_H , where the cross sections become constant. Nevertheless, it becomes a few times as large as the corresponding one obtained by assuming that the "radius" of the nucleon were determined by the meson Compton wave-length. This strongly suggests that the dispersion relation would fail to explain the empirical facts at very high energies. The ambiguity from the experimental data does not seem to upset the above conclusion. The calculational method is outlined in Section 2. In the final section the numerical results and the concluding discussion are given.

§ 2. Derivation of the sum of the total cross sections

We start from the new dispersion relations given by W. Gilbert³⁾

$$(\sigma^+ + \sigma^-)_z = (\sigma^+ + \sigma^-)_{z_0} + \frac{(z^2 - z_0^2)}{\kappa} \left[-\frac{2b^2 G^2}{z^2 z_0^2} - \frac{2}{\pi} \text{P} \int_{z_0}^{\infty} \frac{z dz}{z^2 - z_0^2} \frac{\text{Re } D^e(z)}{z^2 - 1(z^2 - z_0^2)} \right], \quad (1)$$

with

$$b = 1/2\kappa, \quad b^2 G^2 = 4\pi f^2,$$

where z is the laboratory meson energy,** f is the unrationalized pseudovector coupling constant, and κ is the rest mass of the nucleon. σ^+ and σ^- are the total cross sections of the reaction $\pi^+ - p$ and $\pi^- - p$, respectively. The scattering amplitude D^e , even in the meson isotopic indices, is normalized as follows,

$$\text{Im } D^e = \gamma E (\sigma^+ + \sigma^-), \quad (2)$$

* Vavilov's calculation of the cross section is entirely nonsense in this respect (P. V. Vavilov, J.E.T.P. (USSR) 32 (1957), 144).

** Throughout in this paper we use the unit $\hbar = c = \mu = 1$, where μ is the meson rest mass.

where γ is the center-of-mass meson momentum and E is the total center-of-mass energy.

A slight modification to the integrand of the principal integral in (1) enables us to make z' unity. Changing the variable z into $\hat{\xi} = (z^2 - 1)^{1/2}$, we obtain

$$(\sigma^+ + \sigma^-)_{\bar{\xi}} = (\sigma^+ + \sigma^-)_0 - \frac{2b^2 G^2}{\kappa(1 + 1/\hat{\xi}^2)} - \frac{2\hat{\xi}^2}{\pi\kappa} \mathcal{P} \int_0^\infty \frac{d\hat{\xi}}{\hat{\xi}^2(\hat{\xi}^2 - \hat{\xi}^2)} \{ \text{Re } D^e(\hat{\xi}) - \text{Re } D^e(0) \}. \quad (3)$$

We use the same notation $D^e(\hat{\xi})$ to represent a new function $D^e(z = \sqrt{\hat{\xi}^2 + 1})$ of the variable $\hat{\xi}$. The integral is no more singular at $\hat{\xi} = 0$. It is our purpose to check the consistency with the empirical data by calculating the right hand side of (3) at the very high energy $\bar{\xi} = K$.

The main problem in this section is how to express the integral in (3) with the experimental data. Since the empirical phase shifts are available only at low energies (< 300 Mev), we separate the integral region into two sections.

Region A: $0 \leq \hat{\xi} \leq \hat{\xi}_L$. If we choose $\hat{\xi}_L$ being about 3, we can directly evaluate the integral with the experimental phase shifts. The integral in this region is

$$I_A = + \frac{2}{\pi\kappa} \int_0^{\hat{\xi}_L} \frac{d\hat{\xi}}{\hat{\xi}^2(1 - \hat{\xi}^2/K^2)} \{ \text{Re } D^e(\hat{\xi}) - \text{Re } D^e(0) \}. \quad (4)$$

Region B: $\hat{\xi}_L \leq \hat{\xi} < \infty$. In this region we can obtain the real part of the amplitude by means of the usual dispersion relation¹⁾

$$\text{Re } D^e(\hat{\xi}) - \text{Re } D^e(0) = \frac{2b^2 G^2 \hat{\xi}^2}{\hat{\xi}^2 + 1} + \frac{2}{\pi} \kappa \hat{\xi}^2 \mathcal{P} \int_0^\infty \frac{d\hat{\xi}'}{\hat{\xi}'^2 - \hat{\xi}^2} f(\hat{\xi}'), \quad (5)$$

where we abbreviate the sum of the cross sections as

$$f(\hat{\xi}) \equiv \text{Im } D^e(\hat{\xi}) / \kappa \hat{\xi}^2 = (\sigma^+ + \sigma^-)_{\bar{\xi}}. \quad (6)$$

If we substitute (5) for the integrand of the relevant integral, the contribution in this range becomes

$$I_B = - \frac{4b^2 G^2}{\pi\kappa} K^2 \mathcal{P} \int_{\hat{\xi}_L}^\infty \frac{d\hat{\xi}}{(\hat{\xi}^2 - K^2)(\hat{\xi}^2 + 1)} - \frac{4}{\pi^2} K^2 \mathcal{P} \int_{\hat{\xi}_L}^\infty \frac{d\hat{\xi}}{\hat{\xi}^2 - K^2} \mathcal{P} \int_0^\infty \frac{d\hat{\xi}'}{\hat{\xi}'^2 - \hat{\xi}^2} f(\hat{\xi}'). \quad (7)$$

The first integral of (7) is easily calculated and gives

$$C' \equiv - \frac{4b^2 G^2}{\pi\kappa} K^2 \mathcal{P} \int_{\hat{\xi}_L}^\infty \frac{d\hat{\xi}}{(\hat{\xi}^2 - K^2)(\hat{\xi}^2 + 1)} = - \frac{4b^2 G^2}{\pi\kappa} \frac{1}{1 + 1/K^2} \left[\tan^{-1} \frac{\hat{\xi}_L}{K} - \frac{\pi}{2} - \frac{\alpha}{2K} \right]. \quad (8)$$

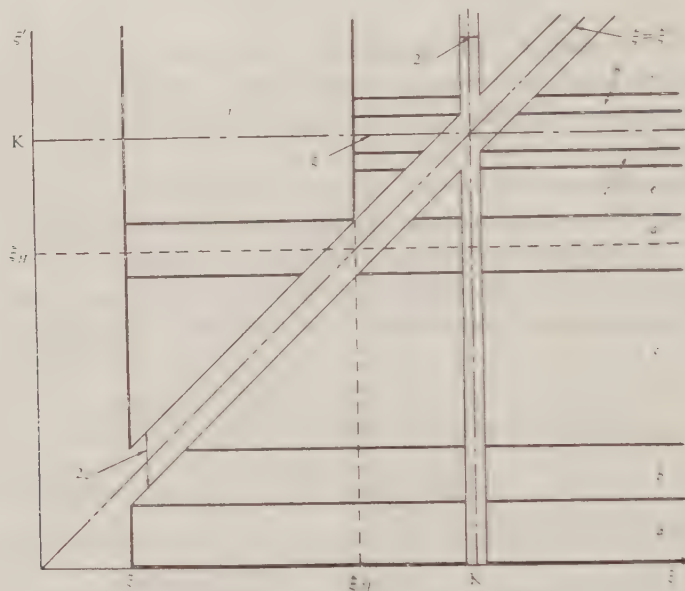
with

$$\alpha \equiv \ln(K - \hat{\xi}_L) / (K + \hat{\xi}_L).$$

The second integral,

$$I_B' = -\frac{4}{\pi^2} K^2 P \int_{\xi_L}^{\infty} \frac{d\xi}{\xi(\xi^2 - K^2)} P \int_{\xi'^2 - \xi}^{\infty} \frac{d\xi'}{\xi'^2 - \xi} f(\xi'), \quad (9)$$

needs a special consideration because of the double principal integration. We separate the region of the double integration into several sections, which are shown in Fig. 1.



- a) ξ ; $(\xi_L \sim K - \varepsilon')$ and $(K + \varepsilon' \sim \infty)$, ξ' ; $(0 \sim \xi_L - \varepsilon)$
- b) ξ ; $(\xi' + \varepsilon \sim K - \varepsilon')$ and $(K + \varepsilon' \sim \infty)$, ξ' ; $(\xi_L - \varepsilon \sim \xi_L + \varepsilon)$
- c) ξ ; $(\xi_L \sim \xi' - \varepsilon)$, $(\xi' + \varepsilon \sim K - \varepsilon')$ and $(K + \varepsilon' \sim \infty)$
 ξ' ; $(\xi_L + \varepsilon \sim \xi_H - \varepsilon)$
- d) ξ ; $(\xi_L \sim \xi' - \varepsilon)$, $(\xi' + \varepsilon \sim K - \varepsilon')$ and $(K + \varepsilon' \sim \infty)$
 ξ' ; $(\xi_H - \varepsilon \sim \xi_H + \varepsilon)$
- e) ξ ; $(\xi_H \sim \xi' - \varepsilon)$, $(\xi' + \varepsilon \sim K - \varepsilon')$ and $(K + \varepsilon' \sim \infty)$
 ξ' ; $(\xi_H + \varepsilon \sim K - \varepsilon - \varepsilon')$
- f) ξ ; $(\xi_H \sim \xi' - \varepsilon)$ and $(K + \varepsilon' \sim \infty)$
 ξ' ; $(K - \varepsilon - \varepsilon' \sim K - \varepsilon + \varepsilon')$
- g) ξ ; $(\xi_H \sim \xi' - \varepsilon)$ and $(\xi' + \varepsilon \sim \infty)$
 ξ' ; $(K - \varepsilon + \varepsilon' \sim K + \varepsilon - \varepsilon')$
- h) ξ ; $(\xi_H \sim K - \varepsilon')$ and $(\xi' + \varepsilon \sim \infty)$
 ξ' ; $(K + \varepsilon - \varepsilon' \sim K + \varepsilon + \varepsilon')$
- i) ξ ; $(\xi_H \sim K - \varepsilon')$, $(K + \varepsilon' \sim \xi' - \varepsilon)$ and $(\xi' + \varepsilon \sim \infty)$
 ξ' ; $(K + \varepsilon + \varepsilon' \sim \infty)$
- j) ξ ; $(\xi_L \sim \xi_H)$, ξ' ; $(\xi_H + \varepsilon \sim \infty)$

Fig. 1. The integral regions of the double integral (9). The small gaps ε and ε' are made to vanish after the integration is carried out. We make $\varepsilon > \varepsilon'$ here, but the result is independent on how to make the limitation.

The principal integration means to make ε and ε' vanish after the integration is performed. Throughout the calculation we always put

$$\varepsilon > \varepsilon' \quad (10)$$

for convenience sake, since we can obtain the unique result independent of how to make the limitation.

After careful calculations, we obtain as the contributions of the integral (9) in each sections

$$\begin{aligned} I'_{Ba} = & -\frac{2}{\pi^2} \left\{ \frac{\alpha}{K} \int_0^{\varepsilon_L} \frac{d\tilde{\varepsilon}' f(\tilde{\varepsilon}')}{1 - \tilde{\varepsilon}'^2/K^2} + \int_0^{\varepsilon_L - \delta} \frac{d\tilde{\varepsilon}' f(\tilde{\varepsilon}')}{\tilde{\varepsilon}' (1 - \tilde{\varepsilon}'^2/K^2)} \ln \frac{\tilde{\varepsilon}_L + \tilde{\varepsilon}'}{\tilde{\varepsilon}_L - \tilde{\varepsilon}'} \right. \\ & \left. + \frac{2f(\tilde{\varepsilon}_L)}{1 - \tilde{\varepsilon}_L^2/K^2} \left[G(1) - G\left(\frac{\tilde{\varepsilon}_L - \delta}{\tilde{\varepsilon}_L}\right) \right] \right\}, \end{aligned} \quad (11a)$$

$$I'_{Bb} = 0, \quad (11b)$$

$$\begin{aligned} I'_{Bc} = & -\frac{2}{\pi^2} \left\{ \frac{\alpha}{K} \int_{\varepsilon_L}^{\varepsilon_H} \frac{d\tilde{\varepsilon}' f(\tilde{\varepsilon}')}{1 - \tilde{\varepsilon}'^2/K^2} + \int_{\varepsilon_L + \delta}^{\varepsilon_H} \frac{d\tilde{\varepsilon}' f(\tilde{\varepsilon}')}{\tilde{\varepsilon}' (1 - \tilde{\varepsilon}'^2/K^2)} \ln \frac{\tilde{\varepsilon}' + \tilde{\varepsilon}_L}{\tilde{\varepsilon}' - \tilde{\varepsilon}_L} \right. \\ & \left. + \frac{2f(\tilde{\varepsilon}_L)}{1 - \tilde{\varepsilon}_L^2/K^2} \left[G(1) - G\left(\frac{\tilde{\varepsilon}_L}{\tilde{\varepsilon}_L + \delta}\right) \right] \right\}, \end{aligned} \quad (11c)$$

$$I'_{Bd} = 0, \quad (11d)$$

$$\begin{aligned} I'_{Be} = & -\frac{f(K)}{\pi^2} \left[\beta \ln \frac{2K}{K + \tilde{\varepsilon}_H} + 2G(1) - 4G\left(\frac{\tilde{\varepsilon}_H}{K}\right) \right. \\ & \left. - F\left(\frac{K - \tilde{\varepsilon}_H}{2K}\right) + F\left(\frac{K + \tilde{\varepsilon}_H}{2K}\right) \right], \end{aligned} \quad (11e)$$

$$I'_{Bf} = 0, \quad (11f)$$

$$I'_{Bg} = +\frac{1}{2} f(K), \quad (11g)$$

$$I'_{Bh} = 0, \quad (11h)$$

$$\begin{aligned} I'_{Bi} = & -\frac{f(K)}{\pi^2} \left[-2G(1) + 4G\left(\frac{\tilde{\varepsilon}_H}{K}\right) + F\left(\frac{K - \tilde{\varepsilon}_H}{2K}\right) - F\left(\frac{K + \tilde{\varepsilon}_H}{2K}\right) \right. \\ & \left. + \frac{1}{2} (\ln(K - \tilde{\varepsilon}_H))^2 - \frac{1}{2} (\ln(K + \tilde{\varepsilon}_H))^2 - \beta \ln 2K \right], \end{aligned} \quad (11i)$$

$$\begin{aligned} I'_{Bj} = & -\frac{f(K)}{\pi^2} \left[-\beta(\beta - \alpha) + \frac{\beta^2}{2} - 4G(1) + 4G\left(\frac{\tilde{\varepsilon}_L}{\tilde{\varepsilon}_H}\right) + F(x_4) - F(x_3) \right. \\ & \left. + F(x_2) - F(x_1) - \frac{1}{2} \alpha^2 + \alpha \ln \frac{K - \tilde{\varepsilon}_L}{K - \tilde{\varepsilon}_H} \right], \end{aligned} \quad (11j)$$

where

$$\beta = \ln \frac{K - \hat{\xi}_H}{K + \hat{\xi}_H}, \quad F(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^2}, \quad G(x) = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)^2},$$

$$x_1 = \frac{K - \hat{\xi}_H}{K + \hat{\xi}_L}, \quad x_2 = \frac{K - \hat{\xi}_H}{K - \hat{\xi}_L}, \quad x_3 = \frac{K + \hat{\xi}_L}{K + \hat{\xi}_H}, \quad x_4 = \frac{K - \hat{\xi}_L}{K + \hat{\xi}_H}.$$

The suffixes a, b, \dots, j correspond to each integral region in Fig. 1. δ in (11a) and (11c) is the small interval of the integration over $\hat{\xi}'$ at $\hat{\xi}' = \hat{\xi}_l$ and it is introduced for the sake of the practical convenience to the calculation. For the integrations in the regions where $\hat{\xi}' \geq \hat{\xi}_H$, we have used assumption I. By using (4), (8) and (11), (3) is now reduced to

$$f(K) = f(0) + I_A + C - \Gamma$$

$$- \frac{2}{\pi^2} \left[\int_0^{\hat{\xi}_L - \delta} \frac{d\hat{\xi} f(\hat{\xi})}{\hat{\xi} (1 - \hat{\xi}^2/K^2)} \ln \frac{\hat{\xi}_L + \hat{\xi}}{\hat{\xi}_L - \hat{\xi}} + \int_{\hat{\xi}_L + \delta}^{\hat{\xi}_H} \frac{d\hat{\xi} f(\hat{\xi})}{\hat{\xi} (1 - \hat{\xi}^2/K^2)} \ln \frac{\hat{\xi} - \hat{\xi}_l}{\hat{\xi} - \hat{\xi}_l} + \alpha \int_0^{\hat{\xi}_H} \frac{d\hat{\xi} f(\hat{\xi})}{1 - \hat{\xi}^2/K^2} \right]$$

$$+ \left[1 - \frac{1}{\pi^2} \left\{ 4G\left(\frac{\hat{\xi}_L}{\hat{\xi}_H}\right) - F_1 + F_2 - F_3 + F_4 + \alpha \ln \frac{K - \hat{\xi}_L}{K + \hat{\xi}_H} - \frac{1}{2} \alpha^2 \right\} \right] f(K), \quad (12)$$

where

$$\Gamma = \frac{4}{\pi^2} \frac{f(\hat{\xi}_l)}{1 - \hat{\xi}_l^2/K^2} \left[\frac{\pi^2}{4} - G\left(\frac{\hat{\xi}_l - \delta}{\hat{\xi}_L}\right) - G\left(\frac{\hat{\xi}_l}{\hat{\xi}_l + \delta}\right) \right].$$

Here we used the fact that $G(1) = \pi^2/8$. There is no singularity in (12) at $K = \hat{\xi}_H$, since the terms containing β have all cancelled out in (12).

It should be noted that (12) must be reduced to the identity

$$f(K) = f(K), \quad (13)$$

if we make $\hat{\xi}_l$ vanish, because it means that the real part $\text{Re}(D^*(\hat{\xi}) - D^*(0))$ is substituted by (5) over all the energy regions. This requirement is indeed satisfied in (12). It must be emphasized that it is necessary to partially substitute (5) in order to obtain a physically meaningful result.

Two integrals in (12) need the cross sections above ~ 15 , say $\hat{\xi}_M$, which corresponds to about 2 Bev of the meson kinetic energy in the laboratory system. Then, in the energy range $\hat{\xi}_M \leq \hat{\xi} \leq \hat{\xi}_H$ we put

$$f(\hat{\xi}) = \frac{1}{\hat{\xi}_H - \hat{\xi}_M} \left[(f(\hat{\xi}_H) - f(\hat{\xi}_M)) \hat{\xi} + \hat{\xi}_H f(\hat{\xi}_M) - \hat{\xi}_M f(\hat{\xi}_H) \right], \quad (14)$$

according to assumption II. Applying (14) to the relevant integrals in (12), we gain the final results,

$$f(K) = f(0) + I_A + C - (H_1 + H_2 + H_3 + \Gamma) - D$$

$$+ \{1 - (A + B)\} f(K), \quad (15)$$

where

$$\begin{aligned}
 H_1 &= \frac{2}{\pi^2} \int_0^{\xi_L^{-6}} \frac{d\hat{\xi} f(\hat{\xi})}{\hat{\xi}(1-\hat{\xi}^2/K^2)} \ln \frac{\hat{\xi}_L + \hat{\xi}}{\hat{\xi}_L - \hat{\xi}}, \quad H_2 = \frac{2}{\pi^2} \int_{\xi_L^{-6}}^{\xi_M} \frac{d\hat{\xi} f(\hat{\xi})}{\hat{\xi}(1-\hat{\xi}^2/K^2)} \ln \frac{\hat{\xi} + \hat{\xi}_L}{\hat{\xi} - \hat{\xi}_L}, \\
 H_3 &= \frac{2}{\pi^2} \frac{\alpha}{K} \int_0^{\xi_M} \frac{d\hat{\xi} f(\hat{\xi})}{1-\hat{\xi}^2/K^2}, \quad D = \frac{f(\hat{\xi}_M)}{\pi^2} \left(\frac{Y\hat{\xi}_H - XK}{\hat{\xi}_H - \hat{\xi}_M} \right), \\
 A &= \frac{1}{\pi^2} \left\{ 4G \left(\frac{\hat{\xi}_L}{\hat{\xi}_H} \right) - F_1 + F_2 - F_3 + F_4 + \alpha \ln \frac{K - \hat{\xi}_L}{K + \hat{\xi}_H} - \frac{1}{2} \alpha^2 \right\}, \\
 B &= \frac{1}{\pi^2} \left(\frac{KX - \hat{\xi}_M Y}{\hat{\xi}_H - \hat{\xi}_M} \right), \\
 X &= (F_1 + F_6 - F_2 - F_5) + (F_4 + F_7 - F_3 - F_8) - \alpha \ln \frac{K + \hat{\xi}_H}{K + \hat{\xi}_M} \\
 Y &= 4G(\hat{\xi}_L, \hat{\xi}_M) - 4G(\hat{\xi}_L, \hat{\xi}_H) + (F_1 + F_6 - F_2 - F_5) \\
 &\quad - (F_4 + F_7 - F_3 - F_8) + \alpha \ln \frac{K + \hat{\xi}_H}{K + \hat{\xi}_M} \\
 x_5 &= \frac{K - \hat{\xi}_M}{K + \hat{\xi}_L}, \quad x_6 = \frac{K - \hat{\xi}_M}{K - \hat{\xi}_L}, \quad x_7 = \frac{K + \hat{\xi}_L}{K + \hat{\xi}_M}, \quad x_8 = \frac{K - \hat{\xi}_L}{K + \hat{\xi}_M},
 \end{aligned} \tag{16}$$

and $F(x_i)$ is abbreviated as F_i .

The first four terms in (15) are completely determined by the experimental data available today. A , B and D are the high energy contributions and depend on our assumptions I and II. B and D represent the contribution from the region $\hat{\xi}_M \leq \hat{\xi} \leq \hat{\xi}_H$, and especially depend on assumption II. Fortunately, these three are, anyhow, very small compared with the other low energy contributions. Then, our conclusions in the next section are fairly insensitive to the assumptions.

§ 3. Numerical results and discussions

The first four terms of (15) can be obtained from the empirical data. $f(0)$, the sum of the cross sections σ^+ and σ^- at the threshold, is determined from the S -wave scattering length by Orear,⁷⁾ which gives

$$f(0) = 0.42. \tag{17}$$

C , the coupling constant term, does not depend on the energy K , which is chosen as $100 \sim \infty$ in the present paper. The result is

$$C = -2.97 f^2 = -0.24, \quad \text{if } f^2 = 0.08. \tag{18}$$

To the integral I_A , which involves the real parts of the scattering amplitudes, the (3, 3) P -wave amplitudes contribute mainly, since the most of the S -wave amplitudes are

subtracted in the integrand and the other P -wave amplitudes are small in magnitude and, further, nearly cancel each other. The result is fairly independent of the choice of K and gives

$$I_A = 6.25. \quad (19)$$

Anderson's empirical formulas presented at the 6th Rochester Conference⁸ are used to obtain (19). The value of I_A is not insensitive to the data at very low energies (see Fig. 2), where there are certain empirical ambiguities. However, the conclusions in this paper are not upset, unless I_A reduces by about 30 percent from the value given in (19) (see Table 1). As is seen in Fig. 2, this is not to be expected.

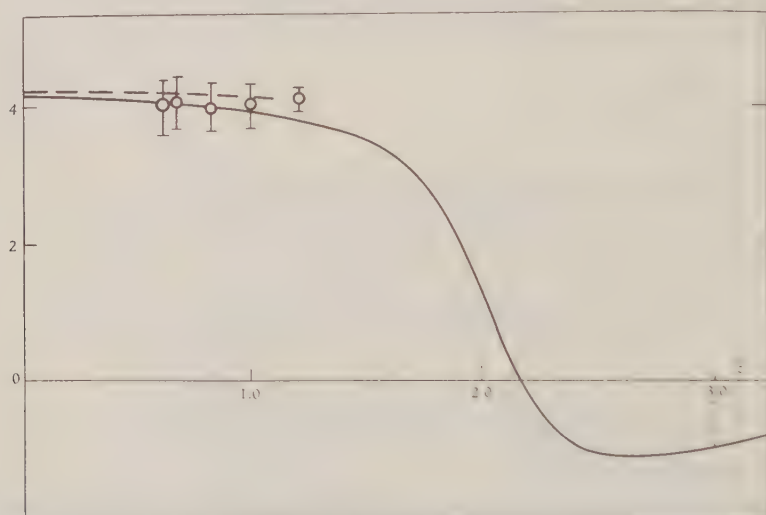


Fig. 2. The integrand of I_A

The solid curve is obtained from Anderson's empirical formulas, and contains the contributions from the small P -wave phase shifts and S -waves. The broken curve is calculated from Chew-Low's static theory, using the formula by G. Puppi et al.⁹⁾ The experimental plots contain only the $(3, 3)$ P -wave phase shift.

H_1 , H_2 , H_3 and I' are calculated using the empirical total cross sections below ~ 2 Bev.^{1), 10)} Each of these and the sum J are given in Table 1 with several choices of $\tilde{\epsilon}_H$ and K . The errors are expected to be less than a few percent as a whole.

A , B and D represent the contributions from high energies ($> \tilde{\epsilon}_H$). These values depend on the choice of $\tilde{\epsilon}_H$, but fairly irrelative to the value of K . The magnitudes are, however, very small compared with the quantities relating to the low energy data. The values of these are also given in Table 1.

Rewriting (15), we see that

$$f(K) = \{ (f(0) + I_A + C) - (J + D) \} / (A + B). \quad (20)$$

Since $(A + B)$ is a very small positive quantity, we must have a very small positive difference,

$$\{ (f(0) + I_A + C) - (J + D) \}, \quad (21)$$

to obtain a reasonable value for $f(K)$, which would be expected to be about 3 (60 mb).

Table 1. The numerical values of the quantities in (15) and (16) with several choices of ξ_H and K . ξ_L and ξ_M are fixed to 3 and 15, respectively.

Case	ξ_H	K	A	B	$A+B$	H_1	H_2
a	15	200	0.0815	—	0.0815	2.76	1.34
b	15		0.0814	—	0.0814	2.75	1.27
c	100	100	0.0122	0.0149	0.0271	2.76	1.48
d	100	200	0.0122	0.0153	0.0275	2.76	1.34
e	100		0.0122	0.0155	0.0277	2.75	1.27

H_3	Γ	$\Delta = \Sigma H_i + \Gamma$	D	$f_0 + I_A + C - (\Delta + D)$	$f(K)$	$(\sigma^+ + \sigma^-)_{mb}$
0.00	0.35	4.45	—	1.98	24.3	486
0.00	0.35	4.37	—	2.06	25.3	506
0.00	0.35	4.59	0.166	1.67	61.5	1230
0.00	0.35	4.45	0.165	1.81	65.8	1320
0.00	0.35	4.37	0.164	1.90	68.6	1370

Unfortunately, it is not the case. In the case $\xi_H=15$, where there is no region to apply assumption II, we obtain about 500 mb as $f(K)$. This is far larger than 60 mb and two times as large as the corresponding geometrical value, which would be obtained if the radius of the nucleon were determined by the meson Compton wavelength. If we choose ξ_H as 100, we have about 1300 mb. The discrepancy is increasing much more in this case. Thus we may conclude that at least one of the fundamental assumptions, under which the dispersion relation holds, might be discarded at very high energies, if the experiments are consistent with the dispersion relations at low energy region.

References

- 1) H. L. Anderson, W. C. Davidon and U. E. Kruse, Phys. Rev. **100** (1955), 339.
G. Puppi and A. Stanghellini, Nuovo Cimento **5** (1957), 1305.
D. Ito, S. Minami and H. Tanaka, Soryushironkenkyu **16** (1958), 591.
T. Tsuchida and A. Kanazawa, Prog. Theor. Phys. **20** (1958), 395.
- 2) M. H. Zaidi and E. L. Lomon, Phys. Rev. **108** (1957), 1352.
T. Tsuchida and A. Kanazawa, Prog. Theor. Phys. **20** (1958), 395.
- 3) W. Gilbert, Phys. Rev. **108** (1957), 1078.
- 4) K. Symanzik, Nuovo Cimento **5** (1957), 659.
A. Arnowitz and G. Feldman, Phys. Rev. **180** (1957), 144.
- 5) R. Cool, O. Piccioni and D. Clark, Phys. Rev. **103** (1956), 1082.
- 6) M. L. Goldberger, Phys. Rev. **99** (1955), 979.
M. L. Goldberger, H. Miyazawa and R. Oehme, Phys. Rev. **99** (1955), 986.
- 7) J. Orear, Phys. Rev. **96** (1954), 176.
- 8) H. L. Anderson, Report on the 6th Rochester Conference.
- 9) G. Puppi and A. Stanghellini, Nuovo Cimento **5** (1957), 1305.

Note added in proof The present method is not so useful to obtain the precise value of the total cross section at infinite energy, because the value of small factor $A+B$ depends sensitively on the assumptions made, i.e., on the choice of ξ_H . It is preferable to say that the extremely large value of $\sigma^+ + \sigma^-$ at very high energies means conversely that the real part of the even amplitude below 300 Mev is not expressed by the integral of the cross sections if $\sigma^+ + \sigma^-$ is kept nearly 60mb at high energies. This just corresponds to Puppi's discrepancy in $\pi^- \rightarrow p$ reaction. In the following paper we shall investigate in detail in what extent the above result is affected by the inaccuracy in the experimental data.

An Analysis of Experimental Data on Extensive Air Showers

Saburo MIYAKE

Institute of Polytechnics, Osaka City University, Osaka

(Received August 4, 1958)

The main interest of the present paper is in analysing the fluctuations in the starting points of extensive air showers in the atmosphere. A method of obtaining the shower curve without taking into account the effect of the fluctuations is presented. The latter effect results in that small energy showers which thrust themselves into the lower atmosphere are observed as if they were large energy showers, and the effect is not insignificant especially in the case of observing showers of large sizes at low altitudes. This may also modify the primary energy spectrum in such a way that the intensity at very high energies is significantly smaller than the one currently adopted. Qualitative interpretations of some unusual events and some speculations on high energy nuclear interactions are also given.

§ 1. Introduction and summary

Up to date, various experiments on extensive air showers have been performed by many authors^{1,2)}. It has been clear that extensive air showers are generated by primary cosmic rays of very high energies. Most of the primaries are protons and the rest are alpha particles and heavier nuclei. Their energies are estimated as high as about $10^{15} - 10^{16}$ eV, and one can obtain some information about such high energy nuclear events only from extensive air showers. An extensive air shower spreads over a wide area whose diameter is as large as several hundreds of meters, and, consequently, this fact facilitates us to observe such high energy events in spite of the very small frequency of the primaries. The information from extensive air showers is important not only for investigating high energy phenomena, but also for studying the mechanism of the retention and acceleration of the particles in our galaxy, through the primary energy spectrum and the relative abundances of heavy nuclei in primary particles. However, since the observation of extensive air showers is not a direct measurement of the high energy events but a partial observation of the great number of particles that have been produced as a result of such high energy events, necessary information is obtained only through the use of statistical analyses.

Recent experimental researches^{1,2)} have shown many phenomenological features of extensive air showers, viz. the size-frequency distribution, the zenith angle distribution, the absorption coefficient in the atmosphere, the lateral distribution of the electron-photon component, μ mesons and the nucleon component at sea level and mountain altitudes, and so on. And the structure of the core of extensive air showers is also being studied at the present time. On the other hand, theoretical calculations³⁾ have been performed about the structure of extensive air showers by applying one of the models of multiple

meson production and electro-magnetic cascade theory. The comparison between such experimental and theoretical results seems to be the only way of studying high energy events.

Although it becomes clear that the extensive air shower is understood as the combination of a high energy nucleon cascade and an electromagnetic cascade which is derived from the decay of π^0 mesons, a unique model that can consistently explain all kinds of the experimental data has not been obtained yet. Moreover, the complexity of the characteristics of the events is superposed on the above difficulty. It must also be mentioned that the results of theoretical calculations so far obtained do not always correspond directly to the observations. Most of theoretical calculations are those of statistical averages over the fluctuations due mainly to those in the atmospheric depth of the first nuclear collision and those further averaged over the primary energy spectrum. However, one can hardly know their primary energies of two such showers with an identical size that are produced possibly from primaries of different energies, if their starting points are different. Therefore, one must be very cautious in comparing the observed data directly with the calculated value.

The purpose of our analysis is to find a clue to studying high energy events as well as to infer the primary energy spectrum important for the origin of cosmic radiation. It is important for this purpose to draw a basic picture from an experimental point of view, not relying only upon theory.

Our main interest is in the fluctuations as has also been pointed out by Kraushaar⁴⁾. Their effects have been studied under probable assumptions on the shower curve; the effects are not at all negligible but, rather, they explain some of the events that have been interpreted as unusual. The characteristic features of extensive air showers without fluctuations in starting points are described on a basis of a different point of view from those so far adopted.

§ 2. Fundamental equation for the observed size distribution

The longitudinal development of an air shower is expressed in terms of the variation of shower size (the total number of particles) with the atmospheric depth. The relation is called "shower curve". A usual shower curve shows an average behaviour of shower phenomena caused by primary particles of a fixed energy, but we shall discuss here a somewhat different implication of the shower curve.

In an air shower, the first stage is a multiple production of secondaries by a collision of a primary cosmic ray particle against an air nucleus and subsequent events are complicated nucleon cascades and accompanied electromagnetic cascades. The size of a shower, once it grows up, does not fluctuate so much on account of the great number of shower particles, but this is not the case in early stages, because the statistical fluctuations in interacting depths are large. Although the fluctuations in later stages might not be negligible, if the inelasticity in a high energy nuclear interaction is small, we may safely neglect their effects, so that we may be allowed to deal only with a fluctuation in the starting points of air showers separately from those in shower developments. After obtaining the

shower curves starting from varying depths, they are summed up with appropriate statistical weights. Thus we obtain a shower curve expressed in terms of simple parameters which are rather directly related to fundamental quantities.

Since the air shower is so complicated that the precise presentation of the shower curve is difficult, the following assumptions are used in the first approximation.

(a) It can be said in the first approximation that the development of an air shower is a function of the air mass through which the air shower has passed after many generations of cascades. Furthermore, the shower curve is a function of the primary energy.

(b) A characteristics of the shower curve is shown as such that the shower size gradually increases with atmospheric depth up to its maximum size and then decreases gradually, but it does not appreciably wiggle. Since, fortunately, the mean free path of nuclear interaction does not much differ from the radiation length of the electromagnetic cascade, the above approximation will be permitted as far as the inelasticity is not too small.

(c) Among the showers which have been generated at a fixed atmospheric depth, a shower by the primary of a larger energy has a larger size than that by the primary of a smaller energy at any depth lower than the generating depth.

(d) The primary cosmic radiation contains protons, alpha particles and heavier nuclei, but the assumptions used here are independent of the nature of these particles; the above properties of the shower curve depends only on the primary energy. Of course the detailed shape of the shower curve depends on the nature of the primaries, but this problem will not be discussed in this paper.

In usual observations, one observes the frequency of air showers which have larger sizes than N at an atmospheric depth X . It can be thought that the observed frequency corresponds to the primary intensity in the following way: since the previous assumptions show that the shower curve is a function of the primary energy E and the depth X , we shall use a threshold C_1 and an observable range C_2 as the representatives of a shower curve, namely, C_1 is the path length necessary to develop the shower up to size N , and C_2 is the range in which the shower is detectable as that with a size exceeding N by usual observations. Then both C_1 and C_2 are functions of the primary energy and the size N :

$$C_1(E, N), C_2(E, N).$$

For the showers whose size is just N at depth X , two limiting cases can be found: one is the shower which starts at the top of the atmosphere and the other is that whose maximum size is equal to N and its position is at depth X . The primary energies of the former and the latter are mentioned as E_2 and E_1 , respectively. (see Fig. 1)

The primary cosmic ray has a mean free path L in the atmosphere in order to generate an air shower effectively, where by "effectively" we mean that the air shower under consideration gives observable effects. L may vary as a function of energy, but, since the variation is thought to be small, L is assumed as constant in the beginning.

According to the above definitions, the frequency of the showers whose sizes are larger

than N at depth X is irrelevant to the intensity of the primaries of energies lower than E_1 but only relevant to the primaries of energies between E_1 and E_2 and larger than E_2 . A shower from a primary energy between E_1 and E_2 is observable when it is started at

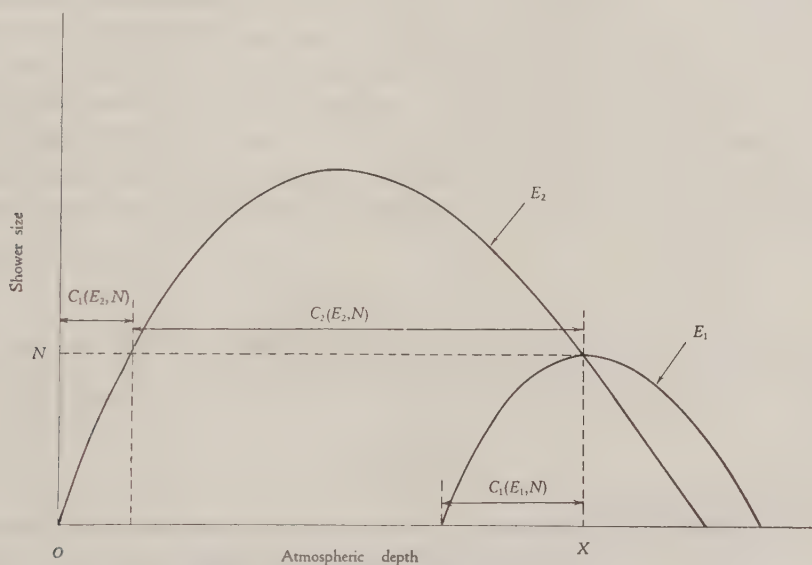


Fig. 1. Shower curves of two limiting cases (E_1 & E_2)

the depth deeper than $X - C_1 - C_2$, and becomes unobservable after the depth $X - C_1$. The probability of detecting such showers is expressed as ;

$$\exp\left(-\frac{X - C_1 - C_2}{L}\right) \cdot \left(1 - \exp\left(-\frac{C_2}{L}\right)\right).$$

The showers initiated by primary energies larger than E_2 are observable when they are started at depths smaller than $X - C_1$, and the probability of detecting them is

$$1 - \exp\left(-\frac{X - C_1}{L}\right).$$

Then the frequency for showers of sizes larger than N is expressed as follows :

$$R(>N) = \int_{E_1}^{E_2} \phi(E) e^{-\frac{(X - C_1 - C_2)}{L}} (1 - e^{-\frac{C_2}{L}}) dE + \int_{E_2}^{\infty} \phi(E) (1 - e^{-\frac{(X - C_1)}{L}}) dE \quad (1)$$

where $\phi(E)dE$ is a differential energy spectrum of primary cosmic radiation.

Quantitative values of the first and the second terms of the right hand side of the above equation must be determined by experiments. We cannot immediately determine the magnitudes of the terms in eq. (1) for the reasons that $\phi(E)$ could be a complicated function of energy ; neither the shower curve nor the ratio of E_1 to E_2 is exactly known. The analysis concerning this problem is described in the next section.

§ 3. Attenuation of the shower size through the atmosphere

Among the experimental results so far obtained, the attenuation of the shower size through the atmosphere has not been fully understood; although a shower of larger size might be regarded to decline more slowly at first sight, the results so far obtained show nearly a constant value of the attenuation length independent of the shower size or a somewhat shorter attenuation length seems to be observed for showers of larger sizes. This seems to be due primarily to the method of estimating the attenuation length. The procedure currently used is such that, first, $R(>N)$ is observed as a function of atmospheric depth (sometimes the effect due to the variation of atmospheric pressure or the zenith angle dependence may be used) and then the absorption coefficient is obtained from the next relation:

$$-\mu = -1/\lambda_{ob} = \partial \ln R / \partial X \quad (2)$$

where λ_{ob} is the absorption (the apparent attenuation) length. Since the observed size distribution is approximately expressed by a power law, if a shower size can be expressed as

$$N = BE^a e^{-X/\Lambda}, \quad (3)$$

the counting rate is given by

$$R(>N) = A(X) N^{-\gamma} = A_0 e^{-X\gamma/\Lambda} N^{-\gamma}. \quad (4)$$

From eq. (2) and eq. (4) one obtains

$$\lambda_{ob} = \Lambda/\gamma. \quad (5)$$

However, there is a question concerning the use of eq. (3); namely we can hardly accept such a simple assumption as eq. (3) from the reason implied in eq. (1). In other words, Λ obtained by eq. (5) represents a complicated quantity which can not always be understood simply as the attenuation length, on account of the fluctuation in the starting points of air showers and the deviation of the shower curve from that expected from this simplified model.

Let us try to apply a method of obtaining an attenuation length similar to what was used for eq. (1). Introducing the following notations for brevity,

$$F_1 = \int_{E_1}^{E_2} \phi(E) e^{-\{X - C_1 - C_2\}/L} (1 - e^{-C_2/L}) dE, \quad (6)$$

$$F_2 = \int_{E_2}^{\infty} \phi(E) dE, \quad (7)$$

$$F_2' = \int_{E_2}^{\infty} \phi(E) (1 - e^{-(X - C_1)/L}) dE, \quad (8)$$

and remembering that C_1 and C_2 are functions of E and N , E_1 is a function of N , E_2 is a function of N and X , and L is a constant, we have

$$-\frac{1}{\lambda_{ob}} = \frac{\partial \ln R}{\partial X} = \frac{-\frac{1}{L} F_1 + \frac{\partial F_1}{\partial E_2} \frac{\partial E_2}{\partial X} - \frac{\partial F_2'}{\partial E_2} \frac{\partial E_2}{\partial X} + \frac{1}{L} (F_2 - F_2')}{F_1 + F_2'} . \quad (9)$$

According to the definition of $E_2(C_1(E_2, N) + C_2(E_2, N) = X)$

$$\frac{\partial F_1}{\partial E_2} \frac{\partial E_2}{\partial X} = \frac{\partial F_2'}{\partial E_2} \frac{\partial E_2}{\partial X} , \quad (10)$$

then eq. (9) is reduced to

$$\lambda_{ob} = L + L \frac{F_2}{F_1 + F_2' - F_2} . \quad (11)$$

By using r defined by

$$r = \frac{F_2}{F_1 + F_2' - F_2} , \quad (12)$$

eq. (11) is expressed as

$$\lambda_{ob} = L + Lr . \quad (13)$$

If $\exp[-(X - C_1)/L]$ is sufficiently small as is expected in the deep atmosphere, F_1 becomes very large compared with $F_2' - F_2$ and r is approximately equal to F_2/F_1 . Hence eq. (13) shows that λ_{ob} does not reveal direct information on the attenuation length; to be more exact, the meaning of λ_{ob} is a quantity which implies both of the mean free path of the primary particles and the penetration of low energy primaries into the atmosphere. The attenuation of the showers is not the one assumed in eq. (3) but it is related to the shower curve in a different way from the current interpretation, as will be described in § 5.

§ 4. The energy spectrum of the primary cosmic radiation

The energy spectrum of the primary cosmic radiation can be obtained from the observed size-frequency distribution by means of a suitable process of transformation. The method of transformation so far used is such that the mean value of the primary energy range, which has been considered to be narrow, is obtained from an observed size and the frequency with this size gives the energy vs intensity relation of the primary cosmic radiation.

Here we will try to derive the primary energy spectrum on the basis of clearer definitions. First, E_2 in eq. (1) was used as the energy which corresponds to the size N owing to the one to one correspondence between energy and size. The value of E_2 can be obtained from the shower curve which will be analysed in § 5; it consists of the energy consumed by the electromagnetic component, which is obtained as a product of the total track length and the critical energy in the air, and the energies of the nucleon component, the μ meson component and the neutrino component which are properly estimated.

As E_2 is a function of X and N , we wish to obtain the relation between E_2 and

N at fixed X , but the observed data are not sufficient for this purpose. A theoretical estimate shows that N is proportional to E^a where a is closed to unity (about 1.1–1.2). In this section, therefore, the energy spectrum is studied under the assumption that E_2 is proportional to N . A modification of this assumption will be described in a later section.

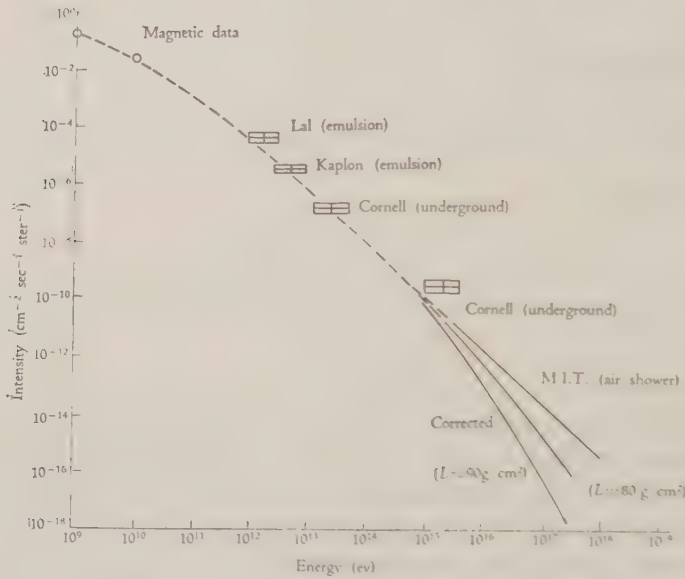


Fig. 2. Integral energy spectrum of primary

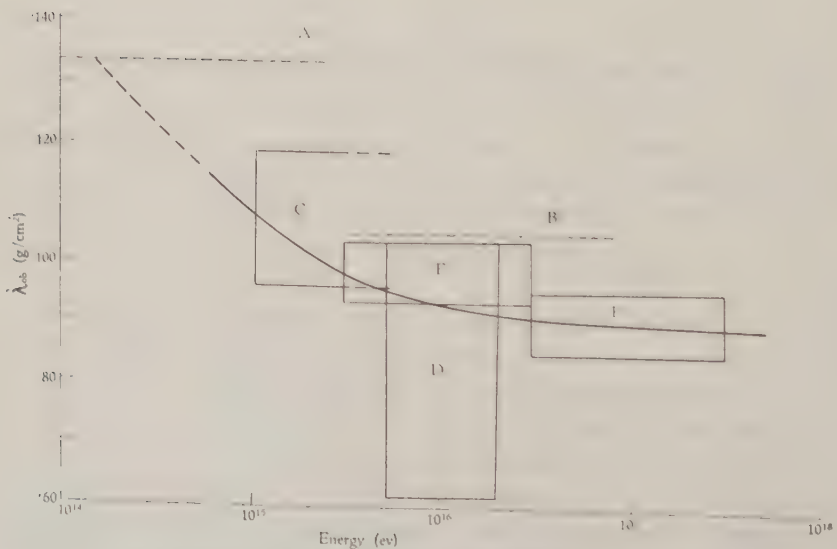


Fig. 3. λ_{ob} versus shower energy (A F: see reference⁵)

The integral frequency at E_2 can be obtained with the aid of the integral frequency at size N , the observed absorption length and the mean free path L ; namely, by eq. (6), (7), (8) and (12) the frequency $\phi(E)dE$ is equal to $R(>N)r/(1+r)$. Since the value of L is not observed yet, two values of L , 80 and 90 g/cm², are adopted as is shown in Fig. 2.

Fig. 2 shows that the intensity of the primary cosmic radiation in a high energy region is quite small in comparison with that adopted in the past. For example, at the size of 3×10^7 (the energy is about 3×10^{17} eV) where λ_{ob} is 90 g/cm², the integral frequency at this energy is negligibly small for $L=90$ g/cm² and it is about one-tenth of the currently adopted intensity for $L=80$ g/cm².

The observed values of λ_{ob} are shown in Fig. 3⁵⁾; it seems to converge to 90 g/cm² as the size increases. If the primaries had an upper limit in their energies, L would be at least 90 g/cm² in the energy region concerned.

The estimates of the primary intensity and L are mutually dependent on one another; moreover, there is a possibility that the relative abundances of the components in the primary radiation vary as the energy increases.

It is desirable to carry out direct measurements of L , which are probably possible by means of the observation of absorption of the showers which have an equal size and an equal shower age at every depth. Another important experiment is the size distribution near the shower maximum, because it approximately reflects the primary energy spectrum.

Since such observations have not been performed yet, we can only conclude, according to the previous simple assumptions, that the primary intensity decreases rapidly in a high energy region and their practical upper limit is at about 5×10^{17} eV. This conclusion is in sharp contrast to that adopted in the past.

§ 5. Shower curve

Observed data available at present are not sufficient to derive the shower curve conclusively. Hence we shall describe only the outline of obtaining the shower curve in this section. The altitude variation of the counting rate has been obtained¹⁾ for a fixed density of shower particles ($\sim 50/\text{m}^2$) and it has also been known already that the rate varies with the density proportional to $J^{-1/5}$. From these facts the shower curve can be obtained by the following way:

(a) First, the observed altitude-frequency curve is reduced to that of the vertical component owing to the relation which has been used by Kraybill and Greisen¹⁾ as follows,

$$F(X) = \frac{C(X)}{2\pi} \left[n+1 - X \frac{\partial \ln C(X)}{\partial X} \right], \quad (14)$$

where $F(X)$ is the integral frequency of the vertical component of the showers which have densities larger than the fixed density J , $C(X)$ is the observed frequency, and n is a constant which, in the usual case of using cylindrical counters, is given by

$$n = 1.8(\gamma - 1) + 0.35\gamma - 0.2 \quad (\gamma = 1.5)$$

where γ is an exponent in the density distribution.

(b) Transformation of the rate $F(>A)$ to the rate $R(>N)$

Assuming that the lateral distribution of the shower particles is inversely proportional to air density and the structure is independent of the atmospheric depth, we express the density at x from the shower axis as

$$A = Nf(x)/x_1 \quad (15)$$

where x_1 is a scattering unit. Since the size distribution has a power dependence on N :

$$R(>N) = R_0 N^{-\tau},$$

we have

$$\begin{aligned} F(>A) &= \int_0^\infty 2\pi x dx R_0 (x_1^2 A/f(x))^{-\tau} \\ &= 2\pi x_1^{-2(\tau-1)} A^{-\tau} R_0 \int (f(x))^\tau x dx. \end{aligned} \quad (16)$$

Accordingly,

$$F(>A) = F_0 A^{-\tau} \propto x_1^{-2(\tau-1)} A^{-\tau} R_0.$$

As x_1 is inversely proportional to air density and the air density is empirically proportional to $X^{0.9}$, the relation of the transformation is as follows:

$$F_0(X) \propto R_0(X) X^{1.8(\tau-1)}. \quad (17)$$

(c) Correction to the effect of the fluctuations

One can obtain λ_{ob} from the gradient of the curve $R(>N)$ versus X at every point of depths and also the value of r by means of eq. (13). Then we can estimate the integral frequency at E_2 by multiplying $R(>N)$ by $r/(1+r)$.

The first term of eq. (1) is absent at altitudes higher than the depth at which $E_1 = E_2$ (this condition takes place at altitudes higher than the depth at which the observed counting rate is maximum). Hence eq. (1) in this region is reduced to

$$R(>N) = \int_{E_2}^\infty \phi(E) dE - \int_{E_2}^\infty \phi(E) e^{-(X-C_1)/L} dE. \quad (18)$$

Let us substitute F_2 and F_3 respectively in the first and the second terms in the right hand side of the above equation then

$$\frac{1}{\lambda_{ob}} = \frac{\partial \ln R}{\partial X} = \frac{-\frac{\partial F_2}{\partial E_2} \frac{\partial E_2}{\partial X} + \frac{1}{L} F_3 + \frac{\partial F_3}{\partial E_2} \frac{\partial E_2}{\partial X}}{F_2 - F_3}. \quad (19)$$

Since $X = C_1$ at the energy of E_2 ,

$$\lambda_{ob} = L \frac{F_2 - F_3}{F_3}. \quad (20)$$

Putting

$$r' = F_3/F_2 = L/(\lambda_{ob} + L), \quad (21)$$

we have

$$\int_{E_2}^{\infty} \phi(E) dE = F_2 = R(>N) \frac{1}{1-r'}. \quad (22)$$

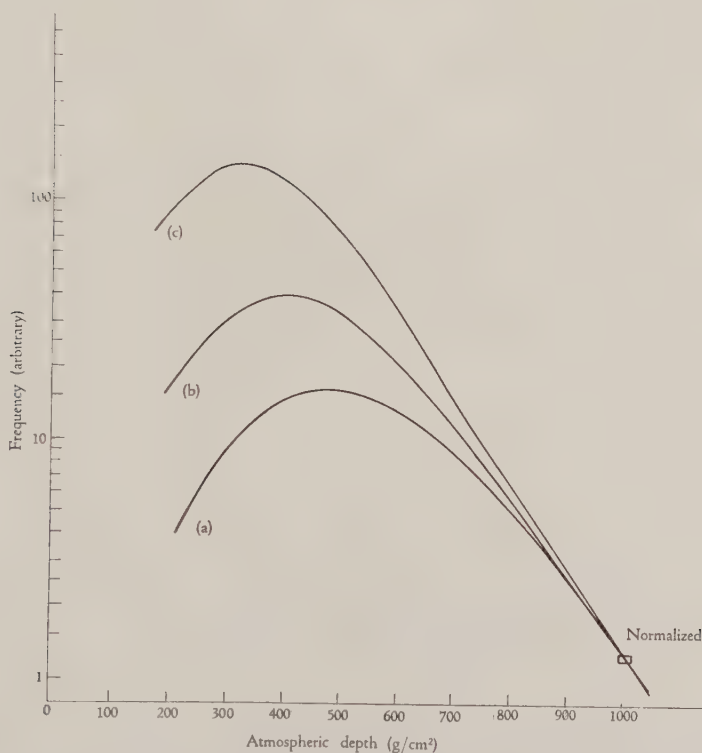


Fig. 4. Frequency-depth relation in § 5, (a) (b) and (c)

(d) Transformation of the frequency distribution into the shower curve

The results of the above paragraphs are shown in Fig. 4 as curves (a), (b) and (c) (these are normalized at the depth of 1 kg/cm² in Fig. 4).

If we have a group of curves which are similar to (c) for every value of N , we can obtain the relation between X and N under the condition that the integral frequency for a fixed energy must be a constant at every depth; namely, the condition means

$$\int_{E_2(X, N)}^{\infty} \phi(E) dE = \text{const}$$

and the shower curve for E_2 indicates the relation between X and N from the above equation.

Unfortunately, the curve is now known only in a narrow range around $N=10^5$ so

that a shower curve is obtained by assuming the size distribution $N^{-1.5}$. The curve is shown in Fig. 5. The position of the maximum is moved step by step by the processes of (a), (b) and (c) and finally it is estimated as about 300 g cm^2 , but its value may be somewhat lower for the primary energy of giving the above shower curve because E_2 at the maximum of the frequency-altitude curve is somewhat lower than the energy under consideration.

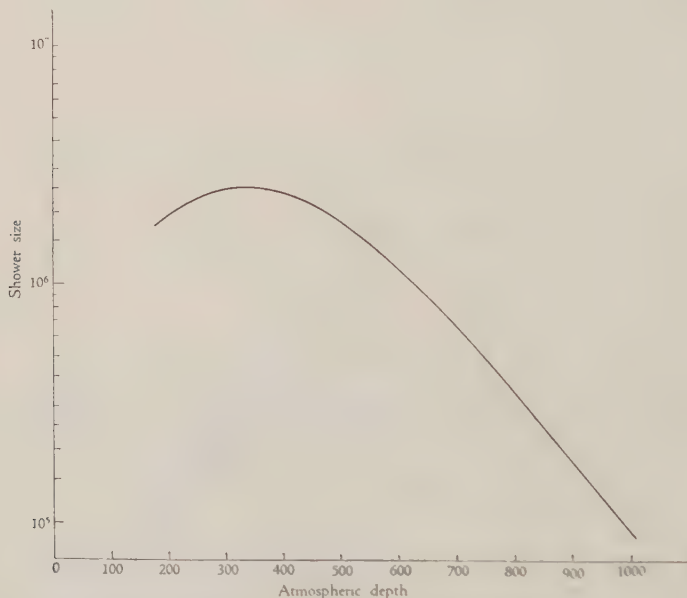


Fig. 5. Shower curve for the energy $\sim 10^{15} \text{ eV}$

§ 6. Discussions

We have pointed out that the treatment of the fluctuations of the air shower development gives a more detailed approximation in the analysis of shower phenomena, but there are so many unknown quantities that a definite conclusion cannot be obtained. When the assumptions used in the previous sections for the characteristics of the air shower were varied, previous discussions would be modified as follows:

(1a) In section 4, the energy spectrum has been obtained under the assumption that the primary energy is proportional to the size at sea level. But the size is thought to be proportional to the power of the primary energy with an exponent which is somewhat larger than unity; accordingly, the tendency that the energy spectrum rapidly decreases in a very high energy region is more strengthened.

(1b) In section 2, the shower curve is assumed to be independent of the altitude of its generation. But in the lower atmosphere, since energetic unstable particles cannot decay owing to the successive collisions with air nuclei, the proportion of the energy which is spent as electromagnetic components becomes somewhat larger and the effect of the fluc-

tuations discussed in the present paper is thought to be greater.

(1c) As described in the last part of section 4, when the value of L and the composition of the primary cosmic radiation, or more generally their characteristics, vary with their energy, the conclusion must be changed. However, this will bring about a very slight modification such that if L varies with energy in a similar manner to that given in Fig. 3 and the primary spectrum is similar to the one currently adopted, their absolute values are somewhat smaller.

The effect of the fluctuations is taken into consideration to explain those phenomena which are hardly understood in terms of the averaged behaviour of air showers. For example,

(2a) The fact that the very energetic nucleon component has been observed⁶⁾ occasionally in the core region of the showers will be explained from the reason that these showers are young; namely, they are showers started at the lower atmosphere.

(2b) The high density μ meson showers observed underground⁷⁾ is also explainable⁸⁾. If such a crude approximation is made that μ mesons are generated from the center of the shower with a fixed angular distribution by a primary particle then the shower which started at the place as low as one-tenth of the average starting altitude will give a large density, about 100 times the usual one.

The statistics of the observation is not yet good enough for the quantitative explanation of such special cases.

Lastly, some of the speculations concerning high energy events will be presented. Of course, it is insufficient to discuss this problem referring only to the shower curve, but one would have to consult with future measurements of the structure of the shower by means of various methods.

(3a) The inelasticity at a collision of a high energy nucleon component against an air nucleus seems to be 0.5–0.7 from the comparison of the shower curve obtained here with the result of a calculation.

(3b) If the largest observed size was equal to a size at the maximum of the showers of an energy near the upper limit of the primary energies, the ratio of the size at the maximum to that at sea level is about 20 to 30. The ratio is nearly equal to that for showers of the primary energy of about 10^{15} eV whose shower curve is obtained in this paper. On the other hand, the attenuation length is thought to be not indifferentiable according to the primary energy. Then it may be concluded that the energy of π^0 mesons which conducts the shower curve does not vary according to the primary energy, and as the energy increases more and more energy goes to heavy mesons and nucleons, etc., than to π mesons.

The author wishes to express his sincere thanks to Professor S. Hayakawa of Kyoto University for his kind encouragement and helpful discussions.

References

- 1) K. Greisen "*Progress in Cosmic Ray Physics*" III (1955) edited by J. G. Wilson, North-Holland Publishing Co. Amsterdam; reports before 1955 are summerized.
- 2) Oxford Conference on the Extensive Air Shower (1956).
- 3) A. Ueda and N. Ogita, *Prog Theor Phys.* **18** (1957), 269.
H. Messel, "*Progress in Cosmic Ray Physics*" II (1954) edited by J. G. Wilson.
M. Oda, *Nuovo Cimento* **5** (1957), 615.
- 4) W. L. Kraushaar, presented in Varenna Conference (1957).
- 5) A: see the reference (1).
B: estimated from the data by R. W. Williams, *Phys. Rev.* **74** (1948), 1689.
C: G. Clark, *Phys. Rev.* **108** (1957), 450.
D: F. Scherb, Private Communication March 4 (1958).
E and F: G. Clark, J. Earl, W. Kraushaar, H. Linsley, B. Rossi and F. Scherb, *Nature* **180** (1957), 353 and **180** (1957), 406.
- 6) N. A. Dobrotin, G. T. Zacepin, S. I. Nikolskij and G. B. Hristiansen, *Sup. Nuovo Cimento* **3** (1956), 635.
S. Miyake, K. Hinotani, I. Katsumata and T. Kaneko, *Jour. Phys. Soc. Japan* **13** (1958), 782.
- 7) S. Higashi, T. Oshio, K. Shibata, K. Watanabe and Y. Watase, *Nuovo Cimento* **5** (1957), 597.
- 8) Y. Sekido, presented at Symposium of Phys. Soc. Japan (1956).

Theory of Unstable Particles in the Wave-Packet-Formalism

Kunio YAMAMOTO*

Department of Physics, Osaka University, Osaka

(Received August 17, 1958)

Scattering problem for finite time interval is discussed by using wave-packet-formalism. A prescription is given which allows us to deal with the observations of different wave packets in an antinomic way. The unstable particle is investigated in connection with the transition probability in the frame of the wave-packet-representation.

§ 1. Introduction

For a recent few years many papers¹⁾⁻⁷⁾ have been published concerning the unitarity of the S -matrix, the renormalization of physical constants, and so on, for the system involving unstable particles. It seems to us quite unsatisfactory, however, that in these papers the unstable character has been formally taken into account, and authors have not touched upon more essential problems such as how to represent a state of unstable particles or how to recognize that an unstable particle actually exists for a finite time interval when we take into consideration the production of unstable particles from stable ones. The difficulties in answering the above problems come mainly from the fact that the unstable state is not the eigenstate of the total Hamiltonian. For instance, Glaser and Källén¹⁾ defined the unstable state in terms of the approximate eigenstate of the total Hamiltonian. The present author²⁾ obtained a similar definition by studying the production and decay of an unstable particle in the stationary treatment. Recently, Nakanishi³⁾ introduced a mathematical notion, complex distribution, and defined the unstable state as one which is nearly identical with an exact eigenstate of the total Hamiltonian with a complex eigenvalue and reached a somewhat different conclusion from the former. Then, he⁴⁾ confirmed his own conclusion by making use of the averaging procedure⁵⁾ for a finite time interval.

In these papers the antinomic character of observations of unstable and those of other states was out of consideration. This character, which is automatically satisfied in the case of stable particles, is also an essential one in the case of unstable particles. This character is usually guaranteed by these two conditions,

$$\sum_i |i\rangle \langle i| = 1 \quad (1.1)$$

$$\text{and} \quad \langle i|j\rangle = \delta_{ij}, \quad (1.2)$$

* Formerly, Kunio NAITO.

that is, realized by choosing a set of states $\{|i\rangle\}$ forming a complete orthonormal set. The condition (1.1) allows us to interpret $|\langle i|\phi\rangle|^2$ as representing the probability for finding antinomic $|i\rangle$ in the state $|\phi\rangle$. (In the theory of probability such a quantity as $|\langle i|\phi\rangle|^2$ is called a "probability distribution".)

The main aim of this paper is to give a prescription for a system involving unstable particles by taking account of the antinomic character mentioned above on the basis of the theory of scattering of wave packet.^{(9),(10)} Consideration on antinomy will enable us to understand the unstable states more clearly and intuitively. As a preliminary discussion, we shall give in § 2 representations of wave packets to interacting particles. In § 3, by using these representations some super-complete set of state vectors will be introduced and the antinomy of these states will be discussed. In the final section, the unstable state will be defined and investigated in connection with the transition probability.

§ 2. Wave-packet-representation

We shall begin with the discussion of the wave packet for the two free particles A and B . The wave packet for the system of A and B having the average momenta \mathbf{p}_1 and \mathbf{p}_2 and the average positions \mathbf{r}_1 and \mathbf{r}_2 respectively at some instant, say, $t=0$, is given by

$$\int d\mathbf{k}_1 d\mathbf{k}_2 F_1(\mathbf{p}_1, \mathbf{r}_1, \mathbf{k}_1) F_2(\mathbf{p}_2, \mathbf{r}_2, \mathbf{k}_2) |k_1, k_2\rangle, \quad (2.1)$$

where
$$F_i(\mathbf{p}, \mathbf{r}, \mathbf{k}) = e^{-i(\mathbf{k}-\mathbf{p})\cdot\mathbf{r}} f_i(\mathbf{k}, \mathbf{p}) \quad i=1, 2 \quad (2.2)$$

is responsible for the shape of the packet and $|k_1, k_2\rangle$ is an eigenvector of the free Hamiltonian for the particles A and B having momenta \mathbf{k}_1 and \mathbf{k}_2 respectively. We assume that $f_i(\mathbf{k}, \mathbf{p})$ is zero except for $\mathbf{k} \approx \mathbf{p}$, and we also assume the normalization

$$\int d\mathbf{k} |f_i(\mathbf{k}, \mathbf{p})|^2 = 1 \quad i=1, 2. \quad (2.3)$$

The wave function of the ordinary space is

$$(2\pi)^{-3/2} e^{i\mathbf{p}\cdot\mathbf{r}} \int d\mathbf{k} e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{r})} f_i(\mathbf{k}, \mathbf{p}). \quad (2.4)$$

Denoting the uncertainty of the momentum and position by $\Delta\mathbf{p}$ and $\Delta\mathbf{r}$ respectively, we have, of course, $\Delta\mathbf{p} \cdot \Delta\mathbf{r} \approx 1$ by uncertainty principle.

We can classify the wave packets of two particles into the following four classes corresponding to relative momentum and position of two particles.

Suppose a wave packet which represents a state of particles A and B having a sufficiently large distance $|\mathbf{r}_1 - \mathbf{r}_2| \gg |\Delta\mathbf{r}|$ at $t=0$. If A and B are approaching to or departing from each other at this instant, then this packet is defined to belong to the class a or b respectively. If the above A and B will pass or have passed by each other in some instant but never come closely, namely their distance can never be smaller than $\Delta\mathbf{r}$, then this

packet is classified into the class *c*. Finally if *A* is present in the vicinity of *B* with the distance $|\mathbf{r}_1 - \mathbf{r}_2| \leq |\mathbf{Jr}|$, this belongs to the class *d*. Here $|\mathbf{Jr}|$ has been defined by

$$|\mathbf{Jr}| = \text{Max}\{|\mathbf{Jr}_1|, |\mathbf{Jr}_2|\} \quad \text{at } t=0.$$

It will be seen that this classification will become ambiguous owing to the increase of $|\mathbf{Jr}|$, for instance, when the wave packet will be deformed before arriving at their closest position. But such a packet need not be considered, as we shall see later, since this packet gives no contribution to the calculation of transition probability.

When two particles *A* and *B* are interacting with each other, the most straightforward way to obtain the statevector corresponding to (2.1) is to replace $|k_1, k_2\rangle$ in (2.1) with an eigenvector of the total Hamiltonian describing a state of two physical particles with momenta k_1 and k_2 but no mutual interaction, but it is quite evident that such an eigenstate cannot exist because of $[\mathbf{p}_i, H] \neq 0$. In the case of $|\mathbf{r}_1 - \mathbf{r}_2| \gg |\mathbf{Jr}|$, we can consider a wave packet representing one particle state for each one in the very neighbourhood of each particle, accordingly the statevector for the system can be given in some meaning as a product of quantities corresponding to such two packets. Let this state be denoted by

$$|\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2\rangle. \quad (2.5)$$

In addition, we define $|\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2\rangle^\pm$ as follows:

$$|\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2\rangle^\pm = \int d\mathbf{k}_1 d\mathbf{k}_2 F_1(\mathbf{p}_1, \mathbf{r}_1, \mathbf{k}_1) F_2(\mathbf{p}_2, \mathbf{r}_2, \mathbf{k}_2) |k_1, k_2\rangle^\pm. \quad (2.6)$$

Here $|k_1, k_2\rangle^+$ or $|k_1, k_2\rangle^-$ is the outgoing or incoming solution of the Schrödinger equation with the incident momenta k_1 and k_2 . Now we shall consider the quantity ${}^+\langle k_1, k_2 | \mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 \rangle$ for a wave packet of the class *a*. For convenience sake, let the outgoing solution be separated into three parts, the incident plane wave, the outgoing spherical wave and the rest which vanishes rapidly for a large distance between colliding particles. The only contribution to ${}^+\langle k_1, k_2 | \mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 \rangle$ in the packet of class *a* may be expected to come from the first part of the outgoing solution provided that $|\mathbf{p}_i|$ and $|\mathbf{r}_1 - \mathbf{r}_2|$ are sufficiently larger than $|\mathbf{Jp}_i|$ and $|\mathbf{Jr}_i|$. Then it seems reasonable to define that (see Appendix B)

$${}^+\langle k_1, k_2 | \mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 \rangle = F_1(\mathbf{p}_1, \mathbf{r}_1, \mathbf{k}_1) F_2(\mathbf{p}_2, \mathbf{r}_2, \mathbf{k}_2) \quad (2.7)$$

for class *a* and that scalar products of $|\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2\rangle$ with outgoing solutions other than $|k_1, k_2\rangle^+$, or with bound states are zero. The same argument can be performed for incoming solutions and a packet of class *b*. Therefore, recalling that the outgoing or incoming solutions together with bound states form a complete set respectively, we get⁽⁹⁾

$$|\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2\rangle = |\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2\rangle^+ \quad (2.8)$$

for a packet of class *a* and

$$|p_1, r_1; p_2, r_2\rangle^- = |p_1, r_1; p_2, r_2\rangle^- \quad (2.9)$$

for a packet of class *b*.

§ 3. Orthogonality and completeness of wave packet

In this section we shall find a set of the statevectors representing wave packets and satisfying the condition (1.1) which is necessary in order that the observations of the different states are antinomic. As an example of this set, we can take $|p_1, r_1; p_2, r_2\rangle^- \} \equiv \mathfrak{S}$ defined by (2.6). From the definition, we can easily see that

$$\begin{aligned} (2\pi)^{-6} \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{p}_1 d\mathbf{p}_2 |p_1, r_1; p_2, r_2\rangle^- \langle p_1, r_1; p_2, r_2| \\ = \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{k}_1 d\mathbf{k}_2 |f_1(\mathbf{k}_1, \mathbf{p}_1)|^2 \cdot |f_2(\mathbf{k}_2, \mathbf{p}_2)|^2 \cdot |\langle \mathbf{k}_1, \mathbf{k}_2 \rangle^-| \langle \mathbf{k}_1, \mathbf{k}_2| \\ = \int d\mathbf{k}_1 d\mathbf{k}_2 |\langle \mathbf{k}_1, \mathbf{k}_2 \rangle^-| \langle \mathbf{k}_1, \mathbf{k}_2|. \end{aligned} \quad (3.1)$$

Here, we have assumed that $|f_i(\mathbf{k}, \mathbf{p})|^2$ is a function of $\mathbf{k}-\mathbf{p}$ and have used the normalization of (2.3). Now for the sake of simplicity, we assume that $|\langle \mathbf{k}_1, \mathbf{k}_2 \rangle^-|$ forms a complete set. Since the extension to more complicated cases is straightforward, we shall restrict ourselves to this simple case. By virtue of the above assumption and (3.1), we see that \mathfrak{S} is one of the sets of the statevectors which satisfies the condition (1.1). As to the condition (1.2) for orthogonality, we have the following relations:

$$\langle p'_1, r'_1; p'_2, r'_2 | p_1, r_1; p_2, r_2 \rangle^- = \begin{cases} = 1 & \text{for } p_i = p'_i, r_i = r'_i \\ \approx 1 & \text{for } |p_i - p'_i| \approx |\mathbf{J}p_i|, |r_i - r'_i| \approx |\mathbf{J}r_i| \\ = 0 & \text{otherwise.} \end{cases} \quad (3.2)$$

Namely, in our case, the orthogonality between any couple of nearly identical statevectors does not hold and norm of each statevector is equal to unity in contrast with the customary formalism where the statevectors are used to be normalized to δ -function. Owing to this peculiar situation \mathfrak{S} is to be called a "super-complete set". The fact that \mathfrak{S} satisfies

$$(2\pi)^{-6} \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{p}_1 d\mathbf{p}_2 |p_1, r_1; p_2, r_2\rangle^- \langle p_1, r_1; p_2, r_2| = 1 \quad (3.3)$$

can be roughly regarded as due to the device that the norm of each vector is contracted in such a way as to cancel out the super-completeness of \mathfrak{S} .

Consider a statevector $|p_1, r_1; p_2, r_2\rangle^-$ belonging to \mathfrak{S} . As already mentioned above, there exists a subset \mathfrak{S}' of elements in \mathfrak{S} , any element of which is not orthogonal to the above specified vector. Let us consider the following expression (3.4) for a state $|\phi\rangle$

$$(2\pi)^{-6} \int d\mathbf{p}_1 d\mathbf{p}_2 d\mathbf{r}_1 d\mathbf{r}_2 | \langle p_1, r_1; p_2, r_2 | \phi \rangle |^2 \quad (3.4)$$

where the domain D of integration includes the above introduced subset \mathcal{S}' and is sufficiently wider than it. In this case we can interpret (3.4) as the probability for finding in $|\psi\rangle$ the states belonging to the domain D , while if D is smaller than \mathcal{S}' , namely it cannot cover \mathcal{S}' , this kind of interpretation is meaningless. For, in the present formalism, statevectors belonging to \mathcal{S}' cannot be distinguished from each other from the observational point of view. In this restricted sense, the quantity $|\langle \mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 | \psi \rangle|^2$ can be called the probability distribution.

§ 4. Transition matrix and unstable states

In this section, we shall investigate the temporal variation of the wave packet in order to understand what the wave packet $|\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2\rangle^-$ represents and to get the expression for transition probabilities. Let us consider a state of the wave packet of the class \mathbf{a} at $t=0$. This state will become

$$e^{-iHt}|\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2\rangle^+ = \int d\mathbf{k}_1 d\mathbf{k}_2 F_1(\mathbf{p}_1, \mathbf{r}_1, \mathbf{k}_1) F_2(\mathbf{p}_2, \mathbf{r}_2, \mathbf{k}_2) \times \exp\{-i(E_A(\mathbf{k}_1) + E_B(\mathbf{k}_2))t\} |\mathbf{k}_1, \mathbf{k}_2\rangle^+ \quad (4.1)$$

at a time t , where $E_i(\mathbf{k}) = \sqrt{\mathbf{k}^2 + m_i^2}$ is the energy of the physical particle i with the mass m_i ($i=A, B$).

We can rewrite (4.1) as follows,

$$\int d\mathbf{k}_1 d\mathbf{k}_2 F_1(\mathbf{p}_1, \mathbf{r}_1 + \mathbf{V}_A(\mathbf{p}_1)t, \mathbf{k}_1) F_2(\mathbf{p}_2, \mathbf{r}_2 + \mathbf{V}_B(\mathbf{p}_2)t, \mathbf{k}_2) \times \exp\{-i(E_A(\mathbf{k}_1) - \mathbf{k}_1 \cdot \mathbf{V}_A(\mathbf{p}_1) + \mathbf{p}_1 \cdot \mathbf{V}_A(\mathbf{p}_1) + E_B(\mathbf{k}_2) - \mathbf{k}_2 \cdot \mathbf{V}_B(\mathbf{p}_2) + \mathbf{p}_2 \cdot \mathbf{V}_B(\mathbf{p}_2))t\} |\mathbf{k}_1, \mathbf{k}_2\rangle^+, \quad (4.2)$$

where $\mathbf{V}_i(\mathbf{p}) = \mathbf{p}/E_i(\mathbf{p})$. It is easily seen that the exponential factor in (4.2) is nearly constant for the variation of \mathbf{k} 's, since the partial derivatives of these factors with respect to \mathbf{k} 's are nearly equal to zero at $\mathbf{k} \approx \mathbf{p}$ and F vanishes for $\mathbf{k} \neq \mathbf{p}$. In general, this exponential factor gives rise to some deformation of wave packets while traveling, but we shall be concerned with the case where this kind of deformation can be neglected. (Low and Chew⁹⁾ showed for the actual elementary particles that except for very slowly moving electrons this deformation is small enough to make the wave-packet-formalism possible for the traveling distance 100 cm.) From (4.2) we see that the wave packets move freely and their velocities are $\mathbf{V}_i(\mathbf{p})$. Since the wave packet (4.1) for $t=0$ belongs to the class \mathbf{a} , there exists a positive time T for which

$$\mathbf{r}_1 + \mathbf{V}_A(\mathbf{p}_1)T = \mathbf{r}_2 + \mathbf{V}_B(\mathbf{p}_2)T. \quad (4.3)$$

As the arguments of F 's for $t > T$ in (4.2) indicate that these F 's belong to the class \mathbf{b} , to such a statevector it seems difficult to give a simple picture as was pointed out in § 2. Therefore (4.1) or (4.2) represents a system of two physical particles having average momenta \mathbf{p}_1 and \mathbf{p}_2 and approaching to each other until the instant T . However, at T

these two particles collide with each other, accordingly (4.1) or (4.2) can be interpreted as representing the state of this system after the collision for $t > T$. The above consideration gives us a way of interpreting such a complicated statevector.

Let us define the transition matrix by

$$\begin{aligned} & \langle \mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4 | e^{-iHt} | \mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 \rangle^+ \\ &= \int d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3 d\mathbf{k}_4 F_3^*(\mathbf{p}_3, \mathbf{r}_3, \mathbf{k}_3) F_4^*(\mathbf{p}_4, \mathbf{r}_4, \mathbf{k}_4) F_1(\mathbf{p}_1, \mathbf{r}_1, \mathbf{k}_1) F_2(\mathbf{p}_2, \mathbf{r}_2, \mathbf{k}_2) \\ & \times \exp\{-i(E_A(\mathbf{k}_3) + E_B(\mathbf{k}_4))t\} \langle \mathbf{k}_3, \mathbf{k}_4 | S | \mathbf{k}_1, \mathbf{k}_2 \rangle, \end{aligned} \quad (4.4)$$

where we have used the definition of the so-called S -matrix $\langle \mathbf{k}_3, \mathbf{k}_4 | S | \mathbf{k}_1, \mathbf{k}_2 \rangle = \langle \mathbf{k}_3, \mathbf{k}_4 | \mathbf{k}_1, \mathbf{k}_2 \rangle^+$. The S -matrix can be decomposed into two terms.

$$\begin{aligned} \langle \mathbf{k}_3, \mathbf{k}_4 | S | \mathbf{k}_1, \mathbf{k}_2 \rangle &= \delta(\mathbf{k}_1 - \mathbf{k}_3) \delta(\mathbf{k}_2 - \mathbf{k}_4) \\ &+ \delta(\mathbf{k}_3 + \mathbf{k}_4 - \mathbf{k}_1 - \mathbf{k}_2) \delta(E_A(\mathbf{k}_3) + E_B(\mathbf{k}_4) - E_A(\mathbf{k}_1) - E_B(\mathbf{k}_2)) \langle \mathbf{k}_3, \mathbf{k}_4 | R | \mathbf{k}_1, \mathbf{k}_2 \rangle. \end{aligned} \quad (4.5)$$

As the first term describes the transition of two particles without mutual interaction, it is of little interest to us.

We shall inquire into the condition for $\mathbf{p}_1, \mathbf{r}_1, \mathbf{p}_2, \mathbf{r}_2$ for getting a non-vanishing transition probability. First, we shall examine the case where the phase of R -matrix is nearly constant for variation of \mathbf{k}_i at the neighbourhood of the fixed \mathbf{p}_i . In this case, the condition for $\mathbf{p}_3, \mathbf{r}_3, \mathbf{p}_4, \mathbf{r}_4$ is that these must satisfy the following relations (4.6) - (4.8) within the error of $\Delta \mathbf{r}_i$ (see Appendix A).

$$\mathbf{r}_3 - \mathbf{r}_1 - V_A(\mathbf{p}_3)t = (V_A(\mathbf{p}_1) - V_A(\mathbf{p}_3))T, \quad (4.6)$$

$$\mathbf{r}_4 - \mathbf{r}_1 - V_B(\mathbf{p}_4)t = (V_A(\mathbf{p}_1) - V_B(\mathbf{p}_4))T, \quad (4.7)$$

$$\mathbf{r}_2 - \mathbf{r}_1 = (V_A(\mathbf{p}_1) - V_B(\mathbf{p}_2))T. \quad (4.8)$$

Here, T is the same one as that introduced before. Therefore in order that the collision takes place, t must be chosen larger than T . (4.6) and (4.7) for $t = T$ show that the set of parameters $(\mathbf{p}_3, \mathbf{p}_4, \mathbf{r}_3, \mathbf{r}_4)$ corresponds to a packet of the class b , accordingly $|\mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4\rangle^-$ represents a system of physical particles having momenta \mathbf{p}_3 and \mathbf{p}_4 and departing from each other. The transition probability for this process is given by

$$(2\pi)^{-6} \int_0^t d\mathbf{p}_3 d\mathbf{p}_4 d\mathbf{r}_3 d\mathbf{r}_4 |\langle \mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4 | e^{-iHt} | \mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 \rangle^+|^2. \quad (4.9)$$

In the second place, we shall treat the case where the R -matrix has a pole of the following type ($R' = \text{const.}$):

$$R = R' \left((\mathbf{k}_3 + \mathbf{k}_4)^2 - (E_A(\mathbf{k}_3) + E_B(\mathbf{k}_4))^2 + \left(M - \frac{i}{2} \Gamma \right)^2 \right)^{-1}. \quad (4.10)$$

Let us assume that $M \gg \Gamma$, $\Gamma \gg |\Delta \mathbf{p}_i|$ and discuss the case

$$(\mathbf{p}_1 + \mathbf{p}_2)^2 - (E_A(\mathbf{p}_1) + E_B(\mathbf{p}_2))^2 \approx (\mathbf{p}_3 + \mathbf{p}_4)^2 - (E_A(\mathbf{p}_3) + E_B(\mathbf{p}_4))^2 \approx M^2, \quad (4.11)$$

since, otherwise the pole of R -matrix will not give any affection to the condition (4.6)–(4.8). In this case, (4.6) and (4.7) are to be replaced by (see Appendix A)

$$\begin{aligned} \mathbf{r}_3 - \mathbf{r}_1 - V_A(\mathbf{p}_3) \left(t - \frac{E_A(\mathbf{p}_3) + E_B(\mathbf{p}_4)}{M\Gamma} \alpha \right) - \frac{\mathbf{p}_3 + \mathbf{p}_4}{M\Gamma} \alpha \\ = (V_A(\mathbf{p}_1) - V_B(\mathbf{p}_3)) T, \end{aligned} \quad (4.12)$$

$$\begin{aligned} \mathbf{r}_4 - \mathbf{r}_1 - V_B(\mathbf{p}_4) \left(t - \frac{E_A(\mathbf{p}_3) + E_B(\mathbf{p}_4)}{M\Gamma} \alpha \right) - \frac{\mathbf{p}_3 + \mathbf{p}_4}{M\Gamma} \alpha \\ = (V_A(\mathbf{p}_1) - V_B(\mathbf{p}_4)) T \end{aligned} \quad (4.13)$$

and (4.7) remains unchanged. Here α is an arbitrary positive number. The transition matrix (4.9) becomes proportional to $\exp(-\alpha)$ in a rough approximation (see Appendix A). Therefore α should be at most of the order of unity. Comparing (4.12) and (4.13) with (4.6) and (4.7), we see that the relations (4.12) and (4.13) can be satisfied for the set of parameters $(\mathbf{p}_3, \mathbf{p}_4, \mathbf{r}_3, \mathbf{r}_4)$ belonging to the class \mathbf{b} if and only if

$$t > T + \frac{E_A(\mathbf{p}_3) + E_B(\mathbf{p}_4)}{M\Gamma} \alpha. \quad (4.14)$$

In this case the statevector $|\mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4\rangle^-$ represents the same state of particles as mentioned before and the quantity $|\mathbf{p}_3 + \mathbf{p}_4| \alpha / (M\Gamma)$ can be regarded as the distance between P and Q as shown in Fig. 1.

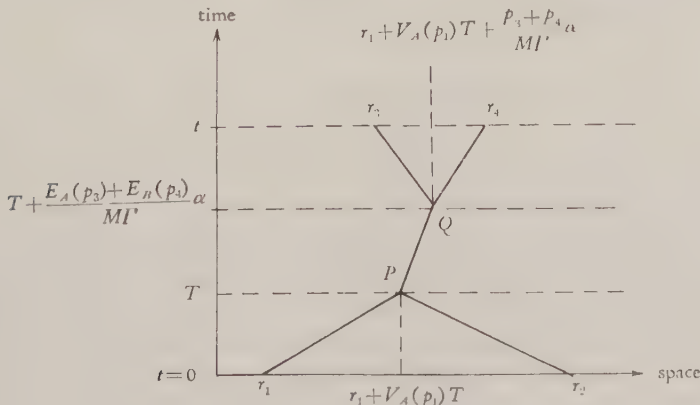


Fig. 1. Schematic representation of the relations (4.12) and (4.13)

The time interval between the world points P and Q can be interpreted as the “life time” during which some dubious state actually exists. The dubious state can exist as an intermediate state with the probability being proportional to $\exp(-\alpha)$. Therefore it

seems quite reasonable to interpret this intermediate state as representing an unstable particle whose velocity is $(\mathbf{p}_3 + \mathbf{p}_4)/(E_A(\mathbf{p}_3) + E_B(\mathbf{p}_4))$ and the mean life time is roughly equal to $1/\Gamma$.

Next, let us consider the case of

$$T + \frac{E_A(\mathbf{p}_3) + E_B(\mathbf{p}_4)}{M\Gamma} \alpha \geq t > T. \quad (4.15)$$

In this case the relations (4.12) and (4.13) can be satisfied only for the set of parameters $(\mathbf{p}_3, \mathbf{p}_4, \mathbf{r}_3, \mathbf{r}_4)$ belonging to the class \mathbf{a} or \mathbf{d} . For such a set of parameters, the statevector $|\mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4\rangle^-$ does not give us a simple picture of the system but we should rather interpret the quantity

$$(2\pi)^{-6} \int_{\mathbf{a} \text{ and } \mathbf{d}} d\mathbf{p}_3 d\mathbf{p}_4 d\mathbf{r}_3 d\mathbf{r}_4 |^-\langle \mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4 | e^{-iHt} | \mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 \rangle|^2 \quad (4.16)$$

as the probability for finding the unstable particle. Accordingly it may be feasible to define unstable state by

$$\int_{\mathbf{a} \text{ and } \mathbf{d}} d\mathbf{p}_3 d\mathbf{p}_4 d\mathbf{r}_3 d\mathbf{r}_4 |\mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4\rangle^- \langle \mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4 | e^{-iHt} | \mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 \rangle^+. \quad (4.17)$$

Now we can rewrite (4.17) as the integral

$$\begin{aligned} & \int_{\mathbf{a} \text{ and } \mathbf{d}} d\mathbf{p}_3 d\mathbf{p}_4 d\mathbf{r}_3 d\mathbf{r}_4 d\mathbf{k}_1 d\mathbf{k}_2 d\mathbf{k}_3 d\mathbf{k}_4 \int_0^\infty d\alpha |\mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4\rangle^- \\ & \times F_3^*(\mathbf{p}_3, \mathbf{r}_3, \mathbf{k}_3) F_4^*(\mathbf{p}_4, \mathbf{r}_4, \mathbf{k}_4) F_1(\mathbf{p}_1, \mathbf{r}_1, \mathbf{k}_1) F_2(\mathbf{p}_2, \mathbf{r}_2, \mathbf{k}_2) \\ & \times \frac{iR'}{M\Gamma} \exp \left[-\alpha - i(E_A(\mathbf{k}_3) + E_B(\mathbf{k}_4))t \right. \\ & \left. - i \left\{ (\mathbf{k}_3 + \mathbf{k}_4)^2 - (E_A(\mathbf{k}_3) + E_B(\mathbf{k}_4))^2 + M^2 - \Gamma^2 \right\} \alpha / (M\Gamma) \right] \end{aligned} \quad (4.18)$$

where the integral parameter α is nothing but the one appearing in (4.12) and (4.13) (see (A.6)). Since the integration with respect to \mathbf{p} 's and \mathbf{r} 's should be made over the classes \mathbf{a} and \mathbf{d} , the relation (4.15) should be satisfied, since otherwise the integrand vanishes. Therefore the domain of integration with respect to α is actually reduced to $(0, (t-T)M\Gamma/(E_A(\mathbf{p}_3) + E_B(\mathbf{p}_4)))$, from which we see that the expression (4.18) decreases exponentially with respect to t , as the function $\exp -i[(\mathbf{k}_3 + \mathbf{k}_4)^2 - (E_A(\mathbf{k}_3) + E_B(\mathbf{k}_4))^2 + M^2 - \Gamma^2]\alpha/(M\Gamma)$ is in practice nearly independent of α (note (4.11)).

As was discussed at the beginning of this section, the expression (4.1) is equal to $|\mathbf{p}_1, \mathbf{r}_1 + \mathbf{V}_A(\mathbf{p}_1)t; \mathbf{p}_2, \mathbf{r}_2 + \mathbf{V}_B(\mathbf{p}_2)t\rangle$, except the trivial constant factor, provided the deformation of wave packet is neglected. Therefore, (4.17) can be rewritten as

$$\int_{\mathbf{a} \text{ and } \mathbf{d}} d\mathbf{p}_3 d\mathbf{p}_4 d\mathbf{r}_3 d\mathbf{r}_4 |\mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4\rangle^- \langle \mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4 | \mathbf{p}_1, \mathbf{r}_1'; \mathbf{p}_2, \mathbf{r}_2' \rangle^+. \quad (4.19)$$

Here, the relation of $\mathbf{p}_1, \mathbf{r}_1', \mathbf{p}_2, \mathbf{r}_2'$ must belong to the class \mathbf{b} owing to the condition $t > T$. It is quite interesting to see that the decay spectrum can be represented by $|\langle \mathbf{p}_3, \mathbf{r}_3; \mathbf{p}_4, \mathbf{r}_4 | \mathbf{p}_1, \mathbf{r}_1'; \mathbf{p}_2, \mathbf{r}_2' \rangle^+|^2$ as shown by (4.19).

Acknowledgements

The author would like to express his sincere thanks to Prof. R. Utiyama for guidance and encouragement. Thanks are also due to Mr. N. Nakanishi for showing his paper to the author before publication and for many valuable discussions, and to Drs. T. Imamura, S. Sunakawa and Mr. T. Kanki for helpful discussions.

Appendix A

In this appendix, we shall investigate the condition for $\mathbf{p}_1, \mathbf{r}_1, \mathbf{p}_2, \mathbf{r}_2$ and $\mathbf{p}_3, \mathbf{r}_3, \mathbf{p}_4, \mathbf{r}_4$ for getting a non-vanishing transition matrix. First, we shall consider the case where the variation of phase of R -matrix can be neglected. In this case, the variation of the phase of the integrand of (4.4) apart from f_i with respect to \mathbf{k}_i is given by

$$\Delta = \Delta \mathbf{k}_3 \cdot \mathbf{r}_3 + \Delta \mathbf{k}_4 \cdot \mathbf{r}_4 - \Delta \mathbf{k}_1 \cdot \mathbf{r}_1 - \Delta \mathbf{k}_2 \cdot \mathbf{r}_2 - \Delta \mathbf{k}_3 \cdot V_A(\mathbf{p}_3)t - \Delta \mathbf{k}_4 \cdot V_B(\mathbf{p}_4)t. \quad (\text{A} \cdot 1)$$

Now, $\Delta \mathbf{k}_i$'s are restricted by the δ -functions for the total energy momentum. These restrictions are

$$\Delta \mathbf{k}_1 + \Delta \mathbf{k}_2 - \Delta \mathbf{k}_3 - \Delta \mathbf{k}_4 = 0 \quad (\text{A} \cdot 2)$$

$$\text{and} \quad \Delta \mathbf{k}_1 \cdot V_A(\mathbf{p}_1) + \Delta \mathbf{k}_2 \cdot V_B(\mathbf{p}_2) - \Delta \mathbf{k}_3 \cdot V_A(\mathbf{p}_3) - \Delta \mathbf{k}_4 \cdot V_B(\mathbf{p}_4) = 0. \quad (\text{A} \cdot 3)$$

In (A.1)–(A.3), $V_A(\mathbf{k}_i)$ and $V_B(\mathbf{k}_i)$ have been replaced with $V_A(\mathbf{p}_i)$ and $V_B(\mathbf{p}_i)$ by virtue of f_i . By using (A.2), $\Delta \mathbf{k}_1$ can be eliminated from (A.1) and (A.3), namely,

$$\Delta = \Delta \mathbf{k}_3 \cdot (\mathbf{r}_3 - \mathbf{r}_1 - V_A(\mathbf{p}_3)t) + \Delta \mathbf{k}_4 \cdot (\mathbf{r}_4 - \mathbf{r}_1 - V_B(\mathbf{p}_4)t) - \Delta \mathbf{k}_2 \cdot (\mathbf{r}_2 - \mathbf{r}_1) \quad (\text{A} \cdot 4)$$

$$\text{and} \quad \Delta \mathbf{k}_3 \cdot (V_A(\mathbf{p}_1) - V_B(\mathbf{p}_3)) + \Delta \mathbf{k}_4 \cdot (V_A(\mathbf{p}_1) - V_B(\mathbf{p}_4)) - \Delta \mathbf{k}_2 \cdot (V_A(\mathbf{p}_1) - V_B(\mathbf{p}_2)) = 0. \quad (\text{A} \cdot 5)$$

In order that (4.4) is not zero, Δ must be nearly equal to zero for the variation of $\Delta \mathbf{k}_i$ restricted by (A.5), when $\Delta \mathbf{k}_i$ is of the order of $\Delta \mathbf{p}_i$. Using the undeterminate multiplier T , we get (4.6)–(4.8).

Secondly, we shall consider the case where R -matrix has the form of (4.10). The expression (4.10) can be rewritten as

$$R = \frac{iR'}{MF} \int_0^\infty d\alpha \exp \left[-\alpha - i \left\{ (\mathbf{k}_3 + \mathbf{k}_4)^2 - (E_A(\mathbf{k}_3) + E_B(\mathbf{k}_4))^2 + M^2 - \Gamma^2 \right\} \alpha / (M\Gamma) \right]. \quad (\text{A} \cdot 6)$$

The variation Δ' of the phase of the R -matrix is

$$\begin{aligned} \Delta' = & -\frac{\alpha}{MI'} \left[(\mathbf{p}_3 + \mathbf{p}_4) \cdot (\Delta \mathbf{k}_3 + \Delta \mathbf{k}_4) \right. \\ & \left. - (E_A(\mathbf{p}_3) + E_B(\mathbf{p}_4)) (V_A(\mathbf{p}_3) \cdot \Delta \mathbf{k}_3 + V_B(\mathbf{p}_4) \cdot \Delta \mathbf{k}_4) \right]. \quad (\text{A} \cdot 7) \end{aligned}$$

Adding this variation Δ' to (A.4), we can obtain (4.12) and (4.13) together with (4.8). The condition (4.11) is derived from the integration with respect to α . From (4.11) and (A.6), we see in a rough approximation that the transition matrix is proportional to $\exp(-\alpha)$.

Appendix B

Using the Lee model,⁽¹¹⁾ we discuss the relationship of wave packets with outgoing or incoming solutions in more detail. When V -particle is unstable, the outgoing or incoming solution of N - θ scattering N , θ_k separately constitutes the complete set and is given by

$$|N\theta_k\rangle^\pm = \left(\int d\mathbf{q} \chi^\pm(\mathbf{q}, \mathbf{k}) a_{\mathbf{q}}^* \psi_{N^*} + \alpha^\pm(\mathbf{k}) \psi_v^* \right) |0\rangle, \quad (\text{B} \cdot 1)$$

$$\text{where} \quad \chi^\pm(\mathbf{q}, \mathbf{k}) = \delta(\mathbf{q} - \mathbf{k}) - g \frac{f(\omega_q)}{\sqrt{\omega_q}} \frac{\alpha^\pm(\mathbf{k})}{\omega_q - \omega_k \mp i\epsilon}, \quad (\text{B} \cdot 2)$$

$$\alpha^\pm(\mathbf{k}) = \frac{gf(\omega_k)}{\sqrt{2\omega_k}} S^\pm(\omega_k + m_N) \quad (\text{B} \cdot 3)$$

$$\text{and} \quad S^\pm(E) = E - m_0 \pm i\epsilon + g^2 \int d\mathbf{k} \frac{f^2(\omega_k)}{2\omega_k} (\omega_k - E + m_N \mp i\epsilon)^{-1}. \quad (\text{B} \cdot 4)$$

Since the bare N - or θ -particle is identical with the physical N - or θ -one, the state corresponding to (2.5) is

$$|\mathbf{p}, \mathbf{r}\rangle = \int d\mathbf{k} F(\mathbf{p}, \mathbf{r}, \mathbf{k}) a^*(\mathbf{k}) \psi_{N^*} |0\rangle. \quad (\text{B} \cdot 5)$$

Here, \mathbf{r} is the relative co-ordinate between N - and θ -particles. Using (B.5), the quantity corresponding to (2.7) is

$$\pm \langle N\theta_k | \mathbf{p}, \mathbf{r} \rangle = F(\mathbf{p}, \mathbf{r}, \mathbf{k}) - g \int d\mathbf{q} \frac{f(\omega_q)}{\sqrt{2\omega_q}} \frac{\alpha^{\pm*}(\mathbf{k})}{\omega_q - \omega_k \pm i\epsilon} F(\mathbf{p}, \mathbf{r}, \mathbf{q}). \quad (\text{B} \cdot 6)$$

The phase variation of the integrand of the second term of (B.6) is

$$\Delta'' = -\Delta \mathbf{q} \cdot \left(\mathbf{r} \mp V(\mathbf{p}) \frac{\epsilon}{(\omega_k - \omega_q)^2 + \epsilon^2} \right). \quad (\text{B} \cdot 7)$$

Therefore, in order that the second term of (B.6) is not zero, \mathbf{r} and $V(\mathbf{p})$ must have the same direction. In this argument, care must be taken, since the variation of the absolute value of the integrand is larger for $\omega_k \approx \omega_q$. We can see, however, that this argument is correct owing to the fact that the sign of the variation of the absolute value

for some specified Δq depends on the sign of $\omega_k - \omega_\eta$. Hence, we can conclude that

$$^+ \langle N \theta_k | \mathbf{p}, \mathbf{r} \rangle = F(\mathbf{p}, \mathbf{r}, \mathbf{k}) \quad (\text{B} \cdot 8)$$

except for the classes **b** and **d**, and

$$^- \langle N \theta_k | \mathbf{p}, \mathbf{r} \rangle = F(\mathbf{p}, \mathbf{r}, \mathbf{k}) \quad (\text{B} \cdot 9)$$

except for the classes **a** and **d**.

References

- 1) H. Araki, Y. Munakata, M. Kawaguchi and T. Goto, Prog. Theor. Phys. **17** (1957), 419.
- 2) T. Okabayashi and S. Sato, Prog. Theor. Phys. **17** (1957), 30.
- 3) V. Glaser and G. Källén, Nuclear Phys. **2** (1957), 706.
- 4) K. Naito, Prog. Theor. Phys. **18** (1957), 200.
- 5) K. Naito, Prog. Theor. Phys. **18** (1957), 614.
- 6) N. Nakanishi, Prog. Theor. Phys. **19** (1958), 607.
- 7) N. Nakanishi, to be published.
- 8) M. Gell-Mann and M. L. Goldberger, Phys. Rev. **91** (1953), 398.
- 9) F. Low and G. F. Chew, *Scattering Theory Seminar (Lecture note at University of Illinois)*, (1953).
- 10) S. Sunakawa, Prog. Theor. Phys. **14** (1955), 175.
- 11) T. D. Lee, Phys. Rev. **95** (1954), 1329.

Quantum Theory in Pseudo-Hilbert Space

Gaku KONISI and Takesi OGIMOTO

Department of Physics, Osaka University, Osaka

(Received July 31, 1958)

A method for treating the Hilbert space with an indefinite metric operator is given. Any transformation of state vectors and operators appearing in physics is generated by a "pseudo-unitary operator". For such transformations two kinds of transformation properties, "covariant" and "contravariant", are considered, just like in the case of the Lorentz transformations in the Minkowski space. It is shown that the difficulty of the Lorentz non-invariance of Gupta's theory in the quantum electrodynamics (reference 2) is due to the ignorance of such transformation properties. Applying our method to the electromagnetic field, we formulate a covariant quantum electrodynamics.

§ 1. Transformation properties of state vectors

Let us start with an ordinary Hilbert space. We introduce a metric operator $\hat{\xi}$ such as

$$\hat{\xi}^* = \hat{\xi} \quad \text{and} \quad \hat{\xi}^{-1} \text{ exist,} \quad (1.1)$$

where $\hat{\xi}^*$ is the Hermitian conjugate of $\hat{\xi}$. $\hat{\xi}^2 = 1$ need not necessarily hold.

As usual, the states of a physical system are represented by vectors of the Hilbert space, whereas we postulate that the probability should be given by

$$\langle x | \hat{\xi} | x \rangle. \quad (1.2)$$

Since the sign of this "pseudo-norm" (1.2) is, in general, indefinite, in order that it may be interpreted as probability, we must put a subsidiary condition on state vectors which assures the positive definiteness of (1.2) at least for the physically allowable states. (In case of the electromagnetic field the positive definiteness is, as well known, guaranteed by virtue of the Lorentz condition.)

We postulate that the metric operator $\hat{\xi}$ is invariant under any transformation considered in physics. Then the operators which keep the pseudo-norms (1.2) invariant must satisfy

$$U^{-1} = \hat{\xi}^{-1} U^* \hat{\xi} \quad (1.3)$$

where U^* is the Hermitian conjugate of U . We call such an operator U a "pseudo-unitary operator", and it will play an important role in the following discussion on the transformation properties.

Let us consider some pseudo-unitary transformation. If a ket vector $|x\rangle$ is transformed under this transformation by a pseudo-unitary operator U as

$$|x\rangle \rightarrow |x'\rangle = U|x\rangle, \quad (1.4a)$$

we call it a “covariant ket” and mark it by a dot symbol attached below as $|x\rangle_.$. The ket $|x\rangle^* = \hat{\xi}|x\rangle_.$ is then transformed as

$$|x\rangle^* \rightarrow |x'\rangle^* = \hat{\xi}U|x\rangle_ = \hat{\xi}U\hat{\xi}^{-1}|x\rangle^*. \quad (1.4b)$$

Each ket vector with such a transformation property is called a “contravariant ket” and denoted by a raised dot symbol. The operator $V = \hat{\xi}U\hat{\xi}^{-1}$ satisfies

$$V^{-1} = \hat{\xi}V^*\hat{\xi}^{-1} \quad (1.5)$$

and, in general, any operator satisfying the relation (1.5) is to be called a “pseudo-unitary operator of the second kind”, while the operator U introduced above the “pseudo-unitary operator of the first kind”.

As the metric operator $\hat{\xi}$ changes $|x\rangle_.$ into $|x\rangle^*$ and $\hat{\xi}^{-1}$ changes $|x\rangle^*$ into $|x\rangle_.$, it will be convenient to denote them as $\hat{\xi}^{**}$ and $\hat{\xi}_{..}$ respectively, i. e.

$$\hat{\xi}^{**}|x\rangle_ = |x\rangle^*, \quad \hat{\xi}_{..}|x\rangle^* = |x\rangle_ . \quad (1.6)$$

Thus $\hat{\xi}^{**}(\hat{\xi}_{..})$ is the operator which raises (lowers) the dot symbol, quite analogously to $g^{\mu\nu}(g_{\mu\nu})$ in the Minkowski space.

The dot symbols can be attached also to the pseudo-unitary operators of the both kinds, U and $\hat{\xi}U\hat{\xi}^{-1}$, as

$$U \equiv U_{.}^* \quad \text{and} \quad \hat{\xi}U\hat{\xi}^{-1} = \hat{\xi}^{**}U_{.}^*\hat{\xi}_{..} \equiv U_{.}^* . \quad (1.7)$$

With these notations (1.4a) and (1.4b) are rewritten in an elegant way:

$$|x'\rangle_ = U_{.}^*|x\rangle_ \quad \text{and} \quad |x'\rangle^* = U_{.}^*|x\rangle^* . \quad (1.8)$$

The bra vectors dual (in the ordinary sense) to $|x\rangle_.$ and $|x\rangle^*$ are transformed as follows:

$$\begin{aligned} {}_.\langle x| &\rightarrow {}_.\langle x'| = {}_.\langle x|(U_{.}^*)^* \equiv {}_.\langle x|U^{\dagger*} , \\ {}^*\langle x| &\rightarrow {}^*\langle x'| = {}^*\langle x|(U_{.}^*)^* \equiv {}^*\langle x|U^{\dagger} . \end{aligned} \quad (1.9)$$

where

$$U^{\dagger*} \equiv (U_{.}^*)^* = (U_{.}^*)^{-1} \quad \text{and} \quad U^{\dagger} \equiv (U_{.}^*)^* = (U_{.}^*)^{-1} . \quad (1.10)$$

Any bra vectors transformed like ${}_.\langle x|$ or ${}^*\langle x|$ are called “covariant” or “contravariant bras” respectively. Evidently, we have

$${}_.\langle x| = {}^*\langle x|\hat{\xi}_{..}, \quad {}^*\langle x| = {}_.\langle x|\hat{\xi}^{**} \quad (1.11)$$

and

$$U^{\dagger} \equiv \hat{\xi}_{..}U^{\dagger*}\hat{\xi}^{**}, \quad U^{\dagger*} \equiv \hat{\xi}^{**}U^{\dagger}\hat{\xi}_{..} . \quad (1.12)$$

We can immediately see that the only inner products which are invariant under the transformation considered are those of a covariant vector with a contravariant one:

$${}^*\langle x|y\rangle_ = {}_.\langle x|y\rangle^* = {}_.\langle x|\hat{\xi}^{**}|y\rangle_ = {}^*\langle x|\hat{\xi}_{..}|y\rangle^* . \quad (1.13)$$

Of course, there may exist vectors that are neither covariant nor contravariant, but they are irrelevant to the covariant description of the theory.

For the sake of convenience, the space composed of those vectors which have the afore mentioned transformation properties is called the "pseudo-Hilbert space".

§ 2. Transformation properties of operators

We now investigate the transformation properties of operators. There are four kinds of operators that are relevant to the covariant formulation of the theory. They are characterized by the following transformation properties, respectively:

- i) Covariant operators : $\mathcal{Q}_{..} \rightarrow \mathcal{Q}'_{..} = U_{.}^{\cdot} \mathcal{Q}_{..} U^{\dagger \cdot}$
- ii) Contravariant operators : $\mathcal{Q}^{..} \rightarrow \mathcal{Q}'^{..} = U_{.}^{\cdot} \mathcal{Q}^{..} U^{\dagger \cdot}$
- iii) Mixed operators of the first kind : $\mathcal{Q}_{.}^{\cdot} \rightarrow \mathcal{Q}'_{.}^{\cdot} = U_{.}^{\cdot} \mathcal{Q}_{.}^{\cdot} U^{\dagger \cdot}$
- iv) Mixed operators of the second kind : $\mathcal{Q}^{\cdot}_{.} \rightarrow \mathcal{Q}'^{\cdot}_{.} = U_{.}^{\cdot} \mathcal{Q}^{\cdot}_{.} U^{\dagger \cdot}$

(2.1)

The dot symbols again indicate the transformation property of the operator. The metric operators $\hat{\varepsilon}^{..}$ and $\hat{\varepsilon}_{..}$ should be a contravariant and a covariant operators, respectively, in order that their dot symbols may be in conformity with the above definitions (2.1). These transformation characters and the pseudo-unitarity of the U -operators give rise to the numerical invariance of the $\hat{\varepsilon}$'s, which is consistent with the postulate in § 1. We can easily show that the above four kinds of operators are mutually related to each other by raising or lowering the dot symbols by means of the $\hat{\varepsilon}$'s, i. e.,

$$\mathcal{Q}_{..} = \hat{\varepsilon}_{..} \mathcal{Q}^{\cdot}_{.}, \quad \mathcal{Q}^{\cdot}_{.} = \hat{\varepsilon}^{..} \mathcal{Q}_{.}^{\cdot}, \quad \text{and so on.}$$

In particular, for the $\hat{\varepsilon}$'s themselves :

$$\hat{\varepsilon}_{..} \hat{\varepsilon}^{..} = \hat{\varepsilon}^{..} \hat{\varepsilon}_{..} = \hat{\varepsilon}^{\cdot}_{.} = \mathbf{I}^{\cdot}, \quad (2.2)$$

Care must be taken in forming a product of operators. Since we shall deal only with the operators whose transformation properties are (2.1), and since other kinds of operators such as $\mathcal{Q}_{....}$, etc., are irrelevant to our formulation as will be easily seen, we must always form a product of operators in such a manner that it also belongs to one of the four classes (2.1). Thus the dots left non-contracted are always located in the both ends of the product and determines the transformation property of this product. For example, we have

$$\begin{aligned} \mathcal{Q}_{1.}^{\cdot} \mathcal{Q}_{2.}^{\cdot} &= \mathcal{Q}_{1..} \mathcal{Q}_{2}^{..} = (\mathcal{Q}_1 \mathcal{Q}_2)^{\cdot}_{.} \\ \mathcal{Q}_1^{..} \mathcal{Q}_{2..} \mathcal{Q}_3^{..} &= \mathcal{Q}_1^{..} \mathcal{Q}_{2.}^{\cdot} \mathcal{Q}_{3.}^{\cdot} = \cdots = (\mathcal{Q}_1 \mathcal{Q}_2 \mathcal{Q}_3)^{..}, \text{ etc.} \end{aligned} \quad (2.3)$$

The above procedure is henceforth referred to as the "rule of proper products". This rule is also applied to products involving vectors, and we are to deal only with such expressions as follows :

$$\begin{aligned}
Q_*|x\rangle_* &= Q_{**}|x\rangle^* = (Q|x\rangle)_*, \\
^*\langle x|Q_* &= \langle x|Q^{**} = ^*(\langle x|Q), \\
^*\langle x|y\rangle_* &= \langle x|y\rangle^*, \\
^*\langle x|Q_*^*|y\rangle_* &= ^*\langle x|Q_{**}|y\rangle^* = \dots
\end{aligned}
\tag{2.4}$$

As is already mentioned, our proper products correspond to the contractions in the tensor calculus, but there is no counterpart of the direct products of tensors (the only exception is the diadic expressions such as $|x\rangle^* \cdot \langle y|$ in which the dot symbols should not be contracted), because quantities with many dots are irrelevant to the present theory as already emphasized in § 1.

The convention of the rule of proper products allows us to omit the annoying dot symbols in formal calculations.

The expectation value of an operator is assumed to be given by the invariant:

$$^*\langle x|Q_*^*|x\rangle_* = ^*\langle x|Q_{**}|x\rangle^* = \dots \tag{2.5}$$

We define the adjoint conjugation of an operator through the complex conjugation of the expectation value, i. e., if

$$\begin{aligned}
(^*\langle x|Q_*^*|x\rangle_*)^* &= ^*\langle x|(Q_*)^\dagger|x\rangle, \\
(^*\langle x|Q_{**}|x\rangle^*)^* &= ^*\langle x|(Q_{**})^\dagger|x\rangle^*, \text{ etc.},
\end{aligned}
\tag{2.6}$$

hold for an arbitrary vector x , $(Q_*)^\dagger$, etc., are called the adjoints of Q_* , etc., respectively. More explicitly the adjoints can be defined in terms of the Hermitian conjugation as follows:

$$\begin{aligned}
(Q_{**})^\dagger &\equiv (Q_{**})^* \\
(Q^{**})^\dagger &\equiv (Q^{**})^* \\
(Q_*)^\dagger &\equiv (Q_*)^* = \hat{\xi}_{**}(Q_*)^*\hat{\xi}^{**} \\
(Q^*)^\dagger &\equiv (Q^*)^* = \hat{\xi}^{**}(Q^*)^*\hat{\xi}_{**}
\end{aligned}
\tag{2.7}$$

(2.7) shows that the present definition of "adjoint" is equivalent to the former one introduced in (1.10). It is easy to show that the adjoint conjugation conserves the transformation property of the operator. Hence we can write $(Q_{**})^\dagger = Q_{**}^\dagger$, $(Q_*)^\dagger = Q_*^\dagger$, etc. From this reason we see that there can exist "self-adjoint" operators which are identical with their own adjoints. Here it must be noticed that the adjoint of a product of operators is, in general, not identical with the product in the reversed order of the adjoints of the individual operators. From the definition (2.7) we have, for example,

$$\begin{aligned}
(Q_{1**}Q_{2**})^\dagger &= ((Q_{1**}Q_{2**})^*)^\dagger = ((Q_{1**}Q_{2**})^*)^* = (Q_{1**}^*Q_{2**}^*)^* \\
&= (Q_{2**})^*(Q_{1**})^* = Q_{2**}^\dagger Q_{1**}^\dagger (\neq Q_{2**}^\dagger Q_{1**}^\dagger).
\end{aligned}$$

Thus the transformation property is, of course, conserved.

The transformation property of an inverse operator, if exists, is, in general, different from that of the original one:

$$(\mathcal{Q}_{..})^{-1} = \mathcal{Q}^{-1*}, \quad (\mathcal{Q}_{..})^{-1} = \mathcal{Q}^{-1*}, \quad (\mathcal{Q}_{..}^*)^{-1} = \mathcal{Q}^{-1*},$$

$$\text{and} \quad (\mathcal{Q}_{..}^*)^{-1} = \mathcal{Q}^{-1*}.$$
(2.8)

§ 3. Eigenvalue problem

The eigenvalue problem of a self-adjoint operator \mathcal{Q} is investigated. The eigenvalue equation to be used here is somewhat different from the ordinary one:

$$\mathcal{Q}_{..}^* |\omega_\alpha\rangle = \omega_\alpha |\omega_\alpha\rangle,$$

$$\mathcal{Q}_{..} |\omega_\alpha\rangle^* = \omega_\alpha |\omega_\alpha\rangle^* = \omega_\alpha \hat{\mathcal{E}}_{..} |\omega_\alpha\rangle^*.$$
(3.1)

The latter equation differs from the ordinary eigenvalue equation of an Hermitian operator $\mathcal{Q}_{..}$, $\mathcal{Q}_{..} |\omega\rangle = \omega |\omega\rangle$.

The equations dual to (3.1) are

$$^*\langle \omega_\alpha | \mathcal{Q}_{..}^* = \omega_\alpha^* ^*\langle \omega_\alpha |, \text{ etc.}$$

From this and (3.1) we have

$$^*\langle \omega_\alpha | \mathcal{Q}_{..}^* |\omega_\beta\rangle = \omega_\alpha^* ^*\langle \omega_\alpha | \omega_\beta\rangle = \omega_\alpha^* \langle \omega_\alpha | \omega_\beta\rangle,$$

$$(\omega_\alpha^* - \omega_\beta) ^*\langle \omega_\alpha | \omega_\beta\rangle = 0,$$

so that

$$\text{if } ^*\langle \omega_\alpha | \omega_\alpha\rangle \neq 0, \quad \omega_\alpha^* = \omega_\alpha \quad (\text{real eigenvalues}),$$

$$\text{and} \quad \text{if } \omega_\alpha^* \neq \omega_\beta, \quad ^*\langle \omega_\alpha | \omega_\beta\rangle = 0 \quad (\text{orthogonality}).$$
(3.2)

These results are the same as those given by S. Gupta and W. Heisenberg in their recent works.¹⁾

We now assume that the eigenvectors of \mathcal{Q} forms a complete set in the Hilbert space (though not orthonormal in the ordinary sense). Then one can take suitable linear combinations of the $|\omega\rangle$'s such that the new set contains no vectors with vanishing pseudo-norms. After normalization we get a complete "orthonormal" set satisfying

$$^*\langle \omega_\alpha | \omega_\beta\rangle = g_{\alpha\beta} = \epsilon_\alpha \delta_{\alpha\beta}, \quad \epsilon_\alpha = +1 \text{ or } -1.$$
(3.3)

The following equalities are immediate consequences of (3.3) and the completeness:

$$\mathbf{1} = \sum_{\alpha,\beta} |\omega_\alpha\rangle g^{\alpha\beta} ^*\langle \omega_\beta | = \sum_{\alpha,\beta} |\omega_\alpha\rangle ^*\langle \omega_\beta | g^{\alpha\beta}, \quad g^{\alpha\beta} = g_{\alpha\beta}$$

$$\hat{\mathcal{E}}^{**} = \sum_{\alpha,\beta} |\omega_\alpha\rangle ^*\langle \omega_\beta | g^{\alpha\beta} ^*\langle \omega_\beta |, \quad \hat{\mathcal{E}}_{..} = \sum_{\alpha,\beta} |\omega_\alpha\rangle g^{\alpha\beta} ^*\langle \omega_\beta |.$$
(3.4)

§ 4. The electromagnetic field

This section is devoted to formulating the covariant quantum electrodynamics by applying our method to the electromagnetic field, and to pointing out that the difficulty

of the Lorentz non-invariance²³⁾ of Gupta's old theory³⁾ is due to the ignorance of the transformation properties.

The electromagnetic four-potential $A_\mu(x)$ is assumed to satisfy the following commutation relation :

$$[A_\mu(x), A_\nu(y)] = -ig_{\mu\nu}D(x-y). \quad (4.1)$$

Then we can immediately see that $A_\mu^*(x)$, the Hermitian conjugate of $A_\mu(x)$, also satisfies the same commutation relation as (4.1) so that $A_\mu(x)$ and $A_\mu^*(x)$ are mutually related by a similarity transformation operator η , namely,

$$A_\mu(x) = \eta^{-1} A_\mu^*(x) \eta, \quad \mu=0, 1, 2, 3. \quad (4.2)$$

Using (4.2) and its Hermitian conjugate equation, we easily get

$$[\eta^{-1} \eta^*, A_\mu(x)] = 0,$$

owing to which η can be chosen to be Hermitian as can be easily seen. This procedure means to give up the usual Hermitian or anti-Hermitian representation of the operators $A_\mu(x)$, and allows us to formulate the whole scheme in a Lorentz-invariant way.

Suppose that the $A_\mu(x)$'s are transformed by an operator U as

$$A_\mu(x) \rightarrow A_\mu'(x) = U A_\mu(x) U^{-1} \quad (4.3)$$

under some physical transformation such as the Lorentz transformations, the gauge transformations, etc. The operator U must satisfy

$$U \eta^{-1} U^* \eta = c$$

where c is a real number, if we postulate that (4.2) is also valid for the same η after the transformation (4.3). By suitably choosing the arbitrary c -number factor in U we have

$$U^{-1} = \pm \eta^{-1} U^* \eta. \quad (4.4)$$

Here we shall consider only the case of the plus sign, which occurs, at least, for the proper Lorentz transformations and the gauge transformations.

Let us now adopt the operator η as the metric operator (1.1). Then, according to the equations (4.2), (4.3) and (4.4), U is a pseudo-unitary operator of the first kind (U, \cdot), whereas $A_\mu(x)$ is a self-adjoint mixed operator of the first kind ($A_\mu(x), \cdot$).

The Fourier expansion of $A_\mu(x), \cdot$ is

$$A_\mu(x), \cdot = \sum_k \frac{1}{\sqrt{2V k^0}} [a_\mu(k), \cdot e^{i k \cdot x} + a_\mu^\dagger(k), \cdot e^{-i k \cdot x}], \quad k^0 = |\mathbf{k}|$$

and the commutation relation of $a_\mu(k), \cdot$ and $a_\mu^\dagger(k), \cdot$ is given by

$$[a_\mu(k), \cdot, a_\nu^\dagger(l), \cdot] = g_{\mu\nu} \delta_{kl} \mathbf{I}, \quad (4.5)$$

or briefly

$$[a., \cdot, a^\dagger., \cdot] = \epsilon \mathbf{I}, \quad \epsilon = +1 \text{ for } \mu=1, 2, 3; \\ -1 \text{ for } \mu=0. \quad (4.5)'$$

We define the "number operator" N_* by

$$N_* \equiv \epsilon a^\dagger_* a_* \quad (4.6)$$

which is apparently self-adjoint but is not non-negative definite. Therefore it can have even complex eigenvalues as we have seen in the preceding section. However, if N_* can be restricted to have non-negative integral eigenvalues, we can still interpret a_* and a^\dagger_* as the annihilation and the creation operators, respectively.

Now, assuming the existence of the "vacuum state" $|0\rangle_*$ with the property^{1),4)}

$$a_*|0\rangle_* = 0, \quad \langle 0|0\rangle_* = 1, \quad (4.7)$$

we can construct a set of eigenvectors of N_* belonging to non-negative integral eigenvalues as follows:

$$\begin{aligned} |n\rangle_* &\equiv (n!)^{-1/2} (a^\dagger_*)^n |0\rangle_*, \quad \langle n| \equiv (n!)^{-1/2} \langle 0| (a_*)^n, \\ N_* |n\rangle_* &= n |n\rangle_*, \quad n = 0, 1, 2, \dots \end{aligned} \quad (4.8)$$

This set is orthonormal in the sense

$$\langle n|m\rangle_* = \epsilon^n \delta_{nm}. \quad (4.9)$$

By virtue of the assumption $a_*|0\rangle_* = 0$ the space \mathfrak{H}_* spanned by the set $\{|n\rangle_*, n = 0, 1, 2, \dots\}$ is an invariant subspace for the operators a_* and a^\dagger_* . Namely, there does not occur any transition between states belonging to \mathfrak{H}_* and to its complement. Therefore, we can confine ourselves to this subspace, assuming that all the physical states belong to \mathfrak{H}_* . The alternative way for this confinement is to adopt the projection operator into \mathfrak{H}_* :

$$P_* \equiv \sum_{n=0}^{\infty} |n\rangle_* \epsilon^n \langle n|. \quad (4.10)$$

It will be easily seen that we can formulate the quantum electrodynamics in a Lorentz-covariant way by following the line of reasoning of Gupta's old theory. For instance, the number of transversal photons can take only values of non-negative integers owing to the assumption $a_*|0\rangle_* = 0$, that is, no difference occurs concerning the transversal components compared with Gupta's theory.

Before concluding this section, let us investigate the relationship between the present method and Gupta's old theory. The latter postulates that the metric operator should satisfy the following commutation relations:

$$\begin{aligned} [A_j(x)_*, \gamma^{**}] &= 0, \quad \text{for } j=1, 2, 3, \\ [A_0(x)_*, \gamma^{**}] &= 0, \end{aligned} \quad (4.11)$$

which make $A_j(x)$ Hermitian and $A_0(x)$ anti-Hermitian, as seen from (4.2). As the commutation relations (4.11) violate the "rule of proper products" mentioned in § 2, they are not U -invariant. The same is true for the Hermitian character of $A_\mu(x)$. Consequently, Gupta's old theory is not U -invariant; namely, it depends on a special Lorentz frame.

§ 5. Conclusion

In the preceding sections we gave a method for treating the indefinite metric space. Our pseudo-Hilbert space with the inner product $\langle x|y \rangle$, forms a so-called "indefinite vector space" recently presented by Gupta.¹⁾ Though Gupta's indefinite vector space is a mathematically larger entity, its mathematical characters are yet little known. Therefore, it will be inconvenient for theoretical investigations. Our pseudo-Hilbert space is, on the contrary, constructed on the Hilbert space which is a well-known concept in mathematics. And we can investigate, for example, the question of convergence with the help of the knowledge of the original Hilbert space.

However, the validity of this advantage is restricted to the case where all the operators under consideration are bounded in the Hilbert space. Actually, this is not always the case. (For instance, the operator of the finite gauge transformation of the electromagnetic field does not satisfy this condition.) These difficulties are closely connected with those in the current field theory. In this paper we confined ourselves to removing the difficulty of Gupta's old theory without wandering out of the Hilbert space.

In passing, we make a remark on the formalism of the Dirac spinors. The conventional γ_4 -method with the Hermitian representation of the γ 's is evidently a wrong one, if the standpoint that the γ -matrices are to be transformed under the Lorentz transformations is equivalent to that in which the wave function ψ is assumed to be changed instead of γ 's. A way out of this inconvenience is realized by a manner quite similar to that given in § 5, by abandoning the Hermitian representation of the γ 's.⁴⁾

Acknowledgements

The authors would like to express their sincere thanks to Prof. R. Utiyama for his kind guidance and encouragement. Thanks are also due to Asst. Prof. S. Takagi and the members of the Utiyama Laboratory for valuable discussions.

References

- 1) S. Gupta, *Canad. Journ. of Phys.* **31** (1957), 961.
W. Heisenberg, *Nuclear Physics* **4** (1957), 532.
- 2) S. Sunakawa, *Prog. of Theor. Phys.* **19** (1958), 221.
- 3) S. Gupta, *Proc. of Phys. Soc.* **63** (1950), 681.
- 4) J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (1954).

Derivation of the Two Nucleon Potential*

Abraham KLEIN

University of Pennsylvania, Philadelphia, Pennsylvania

and

Bruce H. McCORMICK

Radiation Laboratory, University of California, Berkeley, California

(Received August 22, 1958)

A new derivation of the two-nucleon potential proposed recently by the authors is given which avoids the ambiguous features of the scattering formalism employed previously. It is based on the reduction to a Schrödinger equation of a suitable covariant two-nucleon equation. An approximate evaluation of the fourth-order potential highlights the significant departure of the present proposals from past perturbation calculations, but represents a serious overestimate quantitatively of the additional contributions.

§ 1. Introduction

In a preceding note¹⁾ (referred to as *A*) we described in a qualitative manner what we consider to be the soundest method available for obtaining an approximate potential energy function for two nucleons, namely as the adiabatic and low energy limit of the kernel $V(\mathbf{r}, \mathbf{r}'; W)$ of eq. (15) of *A*. The purpose of this paper is twofold. We wish to redevelop the mathematical foundations required for actually deriving the general form of V and to obtain its local interaction limit, in order to justify our previous assertions concerning the latter.

The sections which follow are not arranged in strictly logical order. Section 2 contains a descriptive account of method with the results carried only to the point achieved previously,^{2),3)} whereas Section 3 first contains the definition and derivation of all the required ingredients under the assumption that the meson-meson interaction can be neglected. Finally in Section 4, we proceed to a relatively simple albeit approximate evaluation of the potential suggested in Section 2, the purpose here being merely to exhibit the order of magnitude of the new terms.**

* Supported in part by the U. S. Atomic Energy Commission and by the U. S. Air Force through the Air Research and Development Command.

** A more accurate evaluation and qualitative comparison with experiment has already been carried out by Konuma, Miyazawa, and Otsuki, Prog. Theor. Phys. 19 (1958), 17. Our own evaluation and quantitative comparison with experiment is in preparation.

§ 2. Description of results

Pursuing the method of A , we start with eq. (1) of that note

$$(G^{(1)}G^{(2)})^{-1}\Psi(x_1x_2)=\int d^4x_1'd^4x_2'I(x_1,x_2;x_1',x_2')\Psi(x_1'x_2'). \quad (1)$$

We immediately specialize I to the case where it describes at most the exchange of two mesons, including all self-interactions but omitting non-linear effects of the meson field. We then have

$$I=I_2+I_4, \quad (2)$$

where

$$I_2(x_1x_2;x_1'x_2')=-ig^2\int d^4\hat{\xi}_1d^4\hat{\xi}_1'\Gamma^{(1)}(x_1x_1';\hat{\xi}_1)\Delta(\hat{\xi}_1-\hat{\xi}_1')\Gamma^{(2)}(x_2x_2';\hat{\xi}_1') \quad (3)$$

$$\begin{aligned} I_4(x_1x_2;x_1'x_2') &= (-ig^2)^2/2!\int d^4\hat{\xi}_1\cdots d^4\hat{\xi}_2'\mathcal{J}^{(1)}(x_1x_1';\hat{\xi}_1\hat{\xi}_2) \\ &\quad \times \Delta(\hat{\xi}_1-\hat{\xi}_1')\Delta(\hat{\xi}_2-\hat{\xi}_2')\mathcal{J}^{(2)}(x_2x_2';\hat{\xi}_1'\hat{\xi}_2') \\ &\quad - (-ig^2)^2\int d^4\hat{\xi}_1\cdots d^4\hat{\xi}_2'\langle x_1|\Gamma(\hat{\xi}_1)G\Gamma(\hat{\xi}_2)|x_1'\rangle^{(1)} \\ &\quad \times \Delta(\hat{\xi}_1-\hat{\xi}_1')\Delta(\hat{\xi}_2-\hat{\xi}_2')\langle x_2|\Gamma(\hat{\xi}_1')G\Gamma(\hat{\xi}_2')|x_2'\rangle^{(2)}. \end{aligned} \quad (4)$$

Here $\Gamma^{(i)}$ is the vertex operator for nucleon i , Δ is the Feynman propagator for the physical meson, and $\mathcal{J}^{(i)}$ is the space-time amplitude for the scattering of a meson by a physical nucleon. Equations (3) and (4) are represented diagrammatically in Fig. 1, double lines representing the propagation of physical particles. The term subtracted in (4), as indicated in the diagram is the iteration of I_2 and therefore constitutes a reducible diagram, four-dimensionally speaking. As also indicated in the figure, there then remains crossed diagrams with "single" nucleon propagation in intermediate states plus the sum of all diagrams which cannot be represented so simply. This statement distinguishes the four-point interactions represented by an oval in the diagram from those which are accorded a rectangular representation.

To reduce eq. (1) as actually given to the form of eq. (15) of A would indeed require detailed knowledge of the forms of the functions which occur in eqs. (3) and (4). Suitable forms for this purpose would be the parametric representations which have been so extensively discussed of late.⁴⁾ We shall not pursue the general program here. Instead we shall proceed as expeditiously as possible to the static limit of the resulting interaction. Under these circumstances it is sufficient to replace the single nucleon and meson propagators immediately by the corresponding free propagators with renormalized mass values⁵⁾ and to consider from the beginning that the other quantities occurring have been renormalized. To order g^2 then, the form of eq. (15) of A (in momentum space) can be obtained correctly by carrying out the k_0 and p_0 integrals in the following simple generalization of an equation essentially given by one of the authors,⁶⁾

$$\begin{aligned}
[W - 2E(\mathbf{p})]\varphi(\mathbf{p}) &= A_+^{(1)}(\mathbf{p}) A_+^{(2)}(-\mathbf{p}) (-\lambda^2) (2\pi i)^{-2} \int d^3k dk_0 dp_0 \\
&\times \left\{ \left[\frac{1}{2} W + p_0 - E(\mathbf{p}) \right]^{-1} + \left[\frac{1}{2} W - p_0 - E(\mathbf{p}) \right]^{-1} \right\} \\
&\times (k^2 + \mu^2)^{-1} \Gamma^{(1)}(\mathbf{p}, \mathbf{p} - \mathbf{k}) \Gamma^{(2)}(-\mathbf{p}, -\mathbf{p} + \mathbf{k}) \\
&\times \left\{ \left[\frac{1}{2} W + p_0 - k_0 - E(\mathbf{p} - \mathbf{k}) \right]^{-1} + \left[\frac{1}{2} W - p_0 + k_0 - E(\mathbf{p} - \mathbf{k}) \right]^{-1} \right\} \varphi(\mathbf{p} - \mathbf{k}). \quad (5)
\end{aligned}$$

This differs from equations discussed extensively in Section 3 of reference 6) only in that the full vertex operators $\Gamma^{(i)}$ occur in place of bare vertices. The contribution to the integrations over k_0 and p_0 from the singularities of these functions will again be small corrections to the terms arising from the poles exhibited explicitly.* Indeed as a partial adiabatic limit, it is correct to equate the relative energy dependence of these operators to zero. The resulting kernel will then exhibit the same structure as the second order perturbation theory result, in the notation of A ,

$$v_2(\mathbf{p}, \mathbf{p} - \mathbf{k}; W) = \frac{\lambda \Gamma^{(1)}(\mathbf{p}, \mathbf{p} - \mathbf{k}) \Gamma^{(2)}(-\mathbf{p}, -\mathbf{p} + \mathbf{k})}{\omega(\mathbf{k}) [W - E(\mathbf{p}) - E(\mathbf{p} - \mathbf{k}) - \omega(\mathbf{k})]}, \quad (6)$$

as asserted previously by the authors.²⁾ With slightly altered notation, this is eq. (37) of ref. 2), which is so decisive in obtaining a suitable form of the *fourth-order* potential. For the static vertex operators which occur in (6), we are to write, for example,

$$\Gamma^{(1)}(\mathbf{p}, \mathbf{p} - \mathbf{k}) = i\sigma^{(1)} \cdot \mathbf{k} \tau^{(1)} \rho(k) / 2M, \quad (7)$$

where $\rho(k)$ can be identified with the source function of the fixed source gradient coupling theory.

In order to complete the derivation by extracting the fourth order kernel $v_4^{(2)}$, we again follow the method of ref. 6), which informs us to add to the right-hand side of eq. (5) an expression of the form

$$\begin{aligned}
&A_+^{(1)}(\mathbf{p}) A_+^{(2)}(-\mathbf{p}) (-\lambda^2) (2\pi i)^{-3} \int d^3k_1 d^3k_2 dk_1 dk_2 dp_1 (k_1^2 + \mu^2)^{-1} (k_2^2 + \mu^2)^{-1} \\
&\times \left\{ \left[\frac{1}{2} W + p_0 - E(\mathbf{p}) \right]^{-1} + \left[\frac{1}{2} W - p_0 - E(\mathbf{p}) \right]^{-1} \right\} \left\{ (2!)^{-1} \gamma_0^{(1)} \gamma_0^{(2)} \right. \\
&\times (\mathbf{p}, -\mathbf{k}_1 | \mathcal{J}^{(1)} | \mathbf{p} - \mathbf{k}_1 - \mathbf{k}_2, \mathbf{k}_2) (-\mathbf{p}, \mathbf{k}_1 | \mathcal{J}^{(2)} | -\mathbf{p} + \mathbf{k}_1 + \mathbf{k}_2, -\mathbf{k}_2) \\
&- \Gamma^{(1)}(\mathbf{p}, \mathbf{p} - \mathbf{k}_1) \Gamma^{(2)}(-\mathbf{p}, -\mathbf{p} + \mathbf{k}_1) \left[\frac{1}{2} W + p_0 - k_{01} - E(\mathbf{p} - \mathbf{k}_1) \right]^{-1} \\
&\times \left[\frac{1}{2} W - p_0 + k_{01} - E(\mathbf{p} - \mathbf{k}_1) \right]^{-1} \\
&\times \Gamma^{(1)}(\mathbf{p} - \mathbf{k}_1, \mathbf{p} - \mathbf{k}_1 - \mathbf{k}_2) \Gamma^{(2)}(-\mathbf{p} + \mathbf{k}_1, -\mathbf{p} + \mathbf{k}_1 + \mathbf{k}_2) \left. \right\} \times
\end{aligned}$$

* These corrections as well as those to the meson propagator will give rise to potentials of range 1.3μ and less. See ref. 5) for some of these arguments.

$$\times \left[\left[\frac{1}{2} W - p_0 + k_{01} + k_{02} - E(\mathbf{p} - \mathbf{k}_1 - \mathbf{k}_2) \right]^{-1} + \left[\frac{1}{2} W + p_0 - k_{10} - k_{20} - E(\mathbf{p} - \mathbf{k}_1 - \mathbf{k}_2) \right]^{-1} \right] \varphi(\mathbf{p} - \mathbf{k}_1 - \mathbf{k}_2). \quad (8)$$

This contains only the Fourier transform of the first (reducible) term of eq. (4) from which is subtracted the iterate of $\mathcal{V}^{(2)}$, eq. (6). The appropriate normalization for the operators $\mathcal{J}^{(i)}$ can be inferred from the Born approximation upon comparison with the corresponding equation of ref. 6) and in any case will be given below.

With the statement of eq. (8), the first object of this paper has essentially been achieved. For if we substitute in it for the $\mathcal{J}^{(i)}$ the Born approximations thereto,* to be called $\mathcal{B}^{(i)}$, the resulting kernel yields upon static approximation the fourth order contribution to the BW potential, augmented only by the finite core size effect contained in eq. (7). The potential resulting thus far can be found summarized in eqs. (47) and (48) of ref. 2).

Let us now suppose that the perturbative contributions to eq. (8) just discussed have been removed. What remains has a well-defined static limit. (The iterated term of eq. (8) is now gone.) It is then safe to take the adiabatic limit prior to the performance of the relative energy integrations by remarking that in this limit, for example,

$$\left[\frac{1}{2} W + p_0 - E(\mathbf{p}) + i\epsilon \right]^{-1} + \left[\frac{1}{2} W - p_0 - E(\mathbf{p}) + i\epsilon \right]^{-1} \rightarrow -2\pi i \delta(p_0), \quad (9)$$

where we have remembered, for the first time explicitly, that all contours are to be taken in the Feynman sense. By means of (9) and a corresponding expression for the last curly bracket of eq. (8), we find that under the further neglect of nucleon momenta, we have for instance

$$(\mathbf{p}, p_0; -\mathbf{k}_1, k_{10} | \mathcal{J}^{(1)} | \mathbf{p} - \mathbf{k}_1 - \mathbf{k}_2, p_0 - k_{10} - k_{20}; \mathbf{k}_2, k_{20}) \rightarrow \langle 0, M; -\mathbf{k}_1, k_0 | \mathcal{J}^{(1)} | 0, M; \mathbf{k}_2, k_0 \rangle \equiv (-\mathbf{k}_1 | \mathcal{J}^{(1)}(k_0) | \mathbf{k}_2), \quad (10)$$

where the latter is the scattering amplitude for a virtual meson with indicated energy and momentum by a fixed nucleon. This procedure eventually results in an additional contribution to the potential describing the exchange of two mesons of form**

$$\begin{aligned} \delta v_4(r) = & \frac{1}{2} i (2\pi)^{-7} \int d^3 k_1 d^3 k_2 dk_0 (k_0^2 - \omega_1^2 + i\epsilon)^{-1} (k_0^2 - \omega_2^2 + i\epsilon)^{-1} \\ & \times \exp \left[i(\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r} \right] \left\{ (-\mathbf{k}_1 | \mathcal{J}^{(1)}(-k_0) | -\mathbf{k}_2) (\mathbf{k}_1 | \mathcal{J}^{(2)}(k_0) | \mathbf{k}_2) \right. \\ & \left. - (-\mathbf{k}_1 | \mathcal{B}^{(1)}(-k_0) | -\mathbf{k}_2) (\mathbf{k}_1 | \mathcal{B}^{(2)}(k_0) | \mathbf{k}_2) \right\} \end{aligned} \quad (11)$$

* In the following sense: Free particle propagators, corrected vertex according to eq. (7) and positive energy nucleon intermediate states only (i. e. only P -wave interaction).

** It should also be noted that by way of redefinition, the coupling constants have been incorporated into the scattering amplitudes in contrast to the previous situation.

Equation (11) is equivalent to eq. (51) of ref. 2). The further evaluation can proceed either by the methods described there by the authors, suitably augmented to include S -wave pion-nucleon scattering or, more conveniently, by utilization of the dispersion relations available from fixed source meson theory for the scattering of S and P -wave mesons.^{7,8} The contribution of higher angular momentum states to the static limit of the two-nucleon interaction should be negligible.

To fix the normalization, we shall record the formulae which are to be substituted into eq. (11) for its further evaluation. Here we adopt the notation of Miyazawa,¹⁰ the equations which follow being essentially those given by him. We write

$$\mathcal{J}_{ij}(k_0) = \mathcal{J}_{ij}^{(S)} + \mathcal{J}_{ij}^{(P)}, \quad (12)$$

where the S -waves are represented by the expression

$$\mathcal{J}_{ij}^{(S)}(k_0) = D(k_0) \tau_j \tau_i + E(k_0) \tau_i \tau_j, \quad (13)$$

$$D(k_0) = \frac{2}{3} \pi (a_1 + 2a_3) - \frac{2}{3} (k_0 \cdot \mu) (a_1 - a_3),$$

$$E(k_0) = \frac{2}{3} \pi (a_1 + 2a_3) + \frac{2}{3} (k_0 \cdot \mu) (a_1 - a_3). \quad (14)$$

We have omitted from D and E integrals over the S -wave cross-sections whose contribution can be neglected beside the terms shown. For the P -waves we write

$$(k_1 | \mathcal{J}_{ij}^{(P)}(k_0) | k_2) = A(k_0) \tau_j \tau_i \sigma \cdot k_2 \sigma \cdot k_1$$

$$+ B(k_0) (\tau_j \tau_i \sigma \cdot k_2 \sigma \cdot k_1 + \tau_i \tau_j \sigma \cdot k_1 \sigma \cdot k_2) + C(k_0) \tau_i \tau_j \sigma \cdot k_1 \sigma \cdot k_2, \quad (15)$$

where, including only the Born approximation and contributions from the resonant 33 state,

$$A(k_0) = \left(\frac{f}{\mu} \right)^2 \frac{1}{k_0 - i\epsilon} + \frac{1}{36\pi} \int_0^\infty \frac{dp}{\omega(p)} \sigma_{33}(\omega_p) \left[\frac{9}{\omega(p) - k_0 - i\epsilon} - \frac{1}{\omega(p) + k_0 - i\epsilon} \right], \quad (16)$$

$$B(k_0) = \frac{1}{12\pi} \int_0^\infty \frac{dp}{\omega(p)} \sigma_{33}(\omega_p) \left[\frac{1}{\omega(p) - k_0 - i\epsilon} + \frac{1}{\omega(p) + k_0 - i\epsilon} \right], \quad (17)$$

$$C(k_0) = - \left(\frac{f}{\mu} \right)^2 \frac{1}{k_0 + i\epsilon} + \frac{1}{36\pi} \int_0^\infty \frac{dp}{\omega(p)} \sigma_{33}(\omega_p) \left[\frac{1}{\omega(p) - k_0 - i\epsilon} - \frac{9}{\omega(p) + k_0 - i\epsilon} \right]. \quad (18)$$

The application of these expressions to eq. (11) and the comparison of the resulting potential with experiment will be presented elsewhere.

§ 3. Structure of the interaction in the linear meson approximation

We turn to the derivation of the results described at the start of the previous section. Let $\psi(x)$ be the field operator for the nucleon field and $\phi(\vec{x})$ that for the meson field, spinor and isotopic indices being suppressed as previously. Let the fundamental field

equation satisfied by $\phi(x)$ be*

$$[\gamma p + M - \mathcal{J}(\phi)]\phi = 0, \quad (19)$$

where $\mathcal{J}(\phi)$ is some given function of ϕ . We shall further require the presence of an external source $J(\xi)$ for $\phi(\xi)$ as described by a coupling term

$$\mathcal{L}'(x) = J(x)\phi(x) \quad (20)$$

in the Lagrangian density of the system. Let θ be any operator. Then under the assumption that $J(x)$ vanishes both in the remote past and future, and that vacuum states $|0, (\pm\infty)\rangle$ can therefore be defined at these epochs, we can define

$$\langle\theta\rangle = \langle 0(+\infty) | \theta | 0(-\infty) \rangle / \langle 0(+\infty) | 0(-\infty) \rangle, \quad (21)$$

such that

$$\langle\theta\rangle|_{J=0} = \langle 0 | \theta | 0 \rangle \quad (22)$$

is the usual vacuum expectation value.

The basic formula of the succeeding development is then the statement⁽⁹⁾ $\langle T$ is Wick's time-ordering symbol)

$$\langle T(\phi(\xi)\theta) \rangle = [\langle\phi(\xi)\rangle - i\partial/\partial J(\xi)]\langle\theta\rangle, \quad (23)$$

which, together with the definition

$$\partial\langle\phi(\xi)\rangle/\partial J(\xi')|_{J=0} = i\langle T(\phi(\xi)\phi(\xi')) \rangle = \mathcal{A}(\xi, \xi'), \quad (24)$$

will find repeated application in what follows. Thus, by means of eqs. (19) and (23), we find that the one nucleon Green's function,

$$G(x, y) = i\langle T(\psi(x)\bar{\psi}(y)) \rangle, \quad (25)$$

satisfies the functional-differential equation

$$\{\gamma p + M - \mathcal{J}[\langle\phi\rangle - i\partial/\partial J]\}G = 1. \quad (26)$$

Similarly the two nucleon Green's function

$$G_{12}(x_1, x_2; y_1, y_2) = i^2\langle T(\psi(x_1)\psi(x_2)\bar{\psi}(y_2)\bar{\psi}(y_1)) \rangle, \quad (27)$$

satisfies the equation

$$\{\gamma p + M - \mathcal{J}[\langle\phi\rangle - i\partial/\partial J]\}^{(1)}\{\gamma p + M - \mathcal{J}[\langle\phi\rangle - i\partial/\partial J]\}^{(2)}G_{12} = 1, \quad (28)$$

where here, for example, 1 is the antisymmetric unit matrix,

$$\langle x_1, x_2 | 1 | y_1, y_2 \rangle = \delta(x_1 - y_1)\delta(x_2 - y_2) - \delta(x_1 - y_2)\delta(x_2 - y_1). \quad (29)$$

Ignoring antisymmetrization in what follows, we write for the formal solution of (28)

$$G_{12} = G^{(1)}[\phi - i\partial/\partial J]G^{(2)}[\phi - i\partial/\partial J]|_{J=0}. \quad (30)$$

* The arbitrariness of $\mathcal{J}(\phi)$ will emphasize the generality of the succeeding approach.

Noting that $\partial/\partial J = \Delta \partial/\partial \phi$, it follows that (30) may be rewritten* in the form

$$\begin{aligned} G_{12} &= G^{(1)}[\phi - i\Delta[\phi](\partial/\partial \phi) - i\Delta[\phi](\partial/\partial \chi)]G^{(2)}[\chi - i\Delta[\chi](\partial/\partial \chi)]_{\phi=\chi=0} \\ &= \exp\{-i\Delta[\phi]\partial^2/\partial \phi \partial \chi\} G^{(1)}[\phi - i\Delta[\phi + i\Delta[\phi]\partial/\partial \chi]\partial/\partial \phi] \\ &\quad \times G^{(2)}[\chi - i\Delta[\chi]\partial/\partial \chi]_{\phi=\chi=0}. \end{aligned} \quad (31)$$

The last form of eq. (31) is not exceedingly useful in the general case. In the special instance in which we assume that Δ is independent of ϕ , i. e.,

$$\partial^n \Delta / (\partial \phi)^n = 0, \quad n > 0, \quad (32)$$

which we may term the linear meson approximation, we obtain, however,

$$\begin{aligned} G_{12} &= \exp[-i\Delta \partial^2/\partial \phi \partial \chi] G^{(1)}[\phi - i\Delta \partial/\partial \phi] G^{(2)}[\chi - i\Delta \partial/\partial \chi]_{\phi=\chi=0} \\ &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int d^4 \xi_1 \cdots d^4 \xi_n \Delta(\xi_1 - \xi_n') \cdots \Delta(\xi_n - \xi_n') \\ &\quad \times [\partial^n G^{(1)}[\phi]/\partial \phi(\xi_1) \cdots \partial \phi(\xi_n)]_{\phi=0} [\partial^n G^{(2)}[\chi]/\partial \chi(\xi_1') \cdots \partial \chi(\xi_n')]_{\chi=0}. \end{aligned} \quad (33)$$

A more general analysis which includes non-linear effects will be carried out in a future publication. Here we shall exploit eq. (33), which is more than sufficient for the practical purposes of this paper. The various terms of that equation have a simple interpretation in that the term of order n represents a propagation with the exchange of n mesons, all nucleon self-interaction being included exactly, in principle.

In our previous work,²⁾ the authors used eq. (33) to discuss the two nucleon potential via the intermediary of the S -matrix to which it gives rise. As pointed out there and more forcefully in the preceding note, such a procedure appears fraught with ambiguities. We therefore proceed, in the spirit of this paper, to derive the equation satisfied by G_{12} from the expression (33) which it satisfies. The method is, of course, time worn, the only difference being in the expressions which actually occur here. We write

$$G_{12} = G^{(1)}G^{(2)} + G^{(1)}G^{(2)}I_R G^{(1)}G^{(2)}, \quad (34)$$

where I_R , a sum of both reducible and irreducible interactions, can be read directly from (33). We further suppose I_R to have the form

$$I_R = I + IG^{(1)}G^{(2)}I + \cdots = \sum_{n=0}^{\infty} I(G^{(1)}G^{(2)}I)^n. \quad (35)$$

It follows that G_{12} satisfies the integral equation:

$$G_{12} = G^{(1)}G^{(2)} + G^{(1)}G^{(2)}IG_{12}. \quad (36)$$

The situation has already been illustrated through fourth order in Fig. 1 and eq. (4). Here we need only to record the precise definitions

$$I(x, y; \xi) = \partial G^{-1}[x, y; \phi]/\partial g\phi(\xi)|_{\phi=0}, \quad (37)$$

* The considerations following immediately are discussed in Appendix A of Ref. 2).

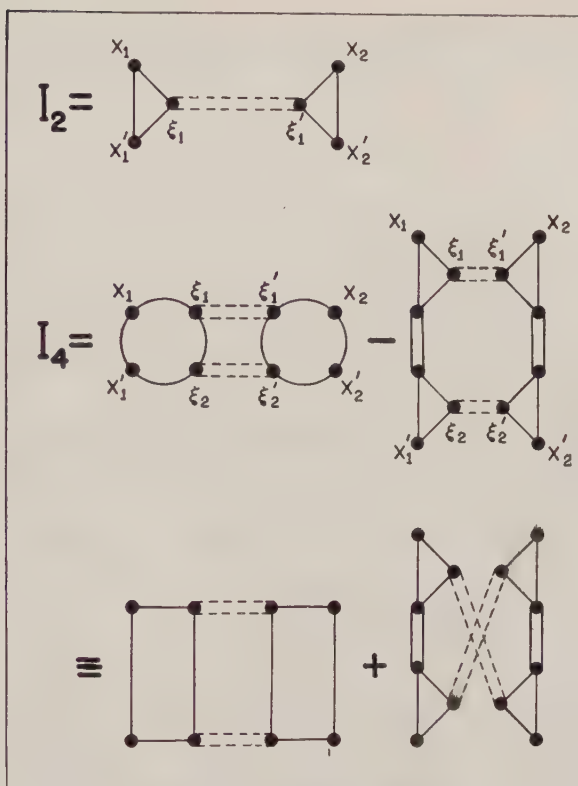


Fig. 1. Feynman diagrams for second and fourth order irreducible interaction. The diagram involving ovals contains reducible terms, whereas that with rectangles contains irreducible interactions only. Double lines indicate the propagation of a physical particle.

and

$$\mathcal{J}(x, \gamma; \xi, \eta) = \langle x | G^{-1} [\partial^2 G / \partial \phi(\xi) \partial \phi(\eta)] G^{-1} | \gamma \rangle_{\phi=0} \quad (38)$$

for the vertex operator and pion-nucleon scattering operator respectively. The results of eq. (4) can easily be extended to all orders in g . Moreover from eq. (36) for G_{12} , there follows eq. (1) for the amplitude $\Psi(x_1, x_2)$,

$$\Psi(x_1, x_2) = \langle 0 | T(\psi(x_1) \psi(x_2)) | \rangle, \quad (39)$$

where $| \rangle$ is an unspecified two nucleon state, the derivation proceeding by means of the limiting procedure of Gell-Mann and Low.¹⁰⁾

§ 4. Approximate evaluation of non-perturbative corrections to fourth order potential*

The importance of the additional potential $\partial v_4(r)$ exhibited in eq. (11) may be

* The results of this section were first reported at the Washington meeting of the American Physical Society, April 1955. See B. H. McCormick and A. Klein, Phys. Rev. **99** (1955), 618.

demonstrated by means of an approach which represents a further simplification of the assumptions of Section 2. We proceed from the appropriate formula for the potential of order $2n$,¹¹⁾

$$v_{2n}(r) \delta(r-r') = -iT^{-1} \int' dt_1 dt_2 dt_1' dt_2' dR' \\ \times \exp[iM(t_1+t_2-t_1'-t_2')] I_{R,2n}(x_1, x_2; x_1', x_2')|_{\text{adiabatic}} \quad (40)$$

which summarizes the experience of this paper. Here T is a "large" time interval which will cancel out in the final result. The prime on the integral signifies that for a reducible part of I_n , eq. (34), the region of four-fold time integration is restricted to exclude times for which there are only two nucleons and no exchange mesons in the intermediate state.

To specialize this result to the calculation of the fourth order potential, we shall utilize $I_{R,4}$ in the form (see eq. (33))

$$I_{R,4}(x_1, x_2; x_1', x_2') = \frac{1}{2} (-i)^2 \int d^4\tilde{z}_1 \cdots d^4\tilde{z}_2 \\ \times \langle x_1 | G^{-1} (\partial^2 G / \partial \phi_{\alpha 1}(\tilde{z}_1) \partial \phi_{\alpha 2}(\tilde{z}_2)) G^{-1} | x_1' \rangle J(\tilde{z}_1 - \tau_{j_1}) J(\tilde{z}_2 - \tau_{j_2}) \\ \times \langle x_2 | G^{-1} (\partial^2 G / \partial \phi_{\alpha 1}(\tau_{j_1}) \partial \phi_{\alpha 2}(\tau_{j_2})) G^{-1} | x_2' \rangle. \quad (41)$$

To evaluate the limit of (41) for fixed nucleons, we require the appropriate second order functional dependence of the single nucleon Green's function G_0 upon an external mesic field ϕ . A consideration of pion-nucleon scattering by one of the authors,¹²⁾ has shown that low energy theorems imply the following approximate non-relativistic form for $G^{-1}[\phi]$.

$$G^{-1}[\phi] = G^{-1}[0] - (g/2M) \int d^4\hat{z} 1(\hat{z}) \sigma_\alpha \tau_i \Gamma_\alpha \phi_i(\hat{z}) \\ + \int d^4\hat{z} d^4\hat{z}' 1(\hat{z}) 1(\hat{z}') \{ \rho_1 (g^2/2M) \phi_i(\hat{z}) \phi_i(\hat{z}') \\ + \rho_2 (g/2M)^2 \epsilon_{ijk} \tau_i \phi_j(\hat{z}) (\partial \phi_k(\hat{z}') / \partial \hat{z}'_0) \\ - \lambda_1 [g^2/(2M)^3] \Gamma_\alpha \phi_i(\hat{z}) \Gamma_\alpha \phi_i(\hat{z}') \\ + \lambda_2 [g^2/(2M)^3] \epsilon_{\alpha\gamma\tau} \epsilon_{ijk} \sigma_\alpha \tau_i \Gamma_\alpha \phi_j(\hat{z}) \Gamma_\tau \phi_k(\hat{z}') \}, \quad (42)$$

where

$$\langle x | 1(\hat{z}) | x' \rangle = \delta^4(x - \hat{z}) \delta^4(x - x'). \quad (43)$$

Equation (42) depends on five parameters: $(g/2M) = (f/\mu)$, the P -wave coupling constant, $(f^2/4\pi) \simeq .08$ two parameters ρ_1 and ρ_2 which measure the strength of the S -wave and a pair of parameters λ_1 and λ_2 which measure the deviation of the P -wave from its Born approximation value. A consequence of eq. (42) is that the S -wave phase shifts, as a function of meson energy ω , are given by expressions of the form

$$\alpha_3(\omega) = - (g^2/4\pi) (\mu/M) \eta [\rho_1 + \frac{1}{2} (\omega/M) \rho_2], \\ \alpha_1(\omega) = - (g^2/4\pi) (\mu/M) \eta [\rho_1 - (\omega/M) \rho_2]. \quad (44)$$

This representation is fully equivalent to that suggested in eqs. (12)–(14) and indeed the present approach treats S -waves with sufficient accuracy for all current needs. Using the experimental values¹³⁾ $\alpha_1/\eta = 0.167 \pm 0.012$, $\alpha_3/\eta = -0.105 \pm 0.010$, we find from (44) (with $\omega = \mu$) that $\rho_1 \cong .01$, $\rho_2 \cong .57$.

Similarly the P -wave phase shifts may be written as

$$\begin{aligned}\alpha_{33}(\omega) &= (4/3) (f^2/4\pi) \eta^3 (\mu/\omega) [1 + \frac{1}{4} (\omega/M) (\lambda_1 + \lambda_2)], \\ \alpha_{31} &= \alpha_{13} = (-2/3) (f^2/4\pi) \eta^3 (\mu/\omega) [1 - \frac{1}{2} (\omega/M) (\lambda_1 - 2\lambda_2)], \\ \alpha_{11} &= -(8/3) (f^2/4\pi) \eta^3 (\mu/\omega) [1 - \frac{1}{8} (\omega/M) (\lambda_1 + 4\lambda_2)].\end{aligned}\quad (45)$$

The representation (45) is equivalent near threshold to the effective range approximation.¹⁴⁾ With $\omega = \mu$ and using the accepted values¹⁵⁾ $\alpha_{33} = 0.22 \eta^3$, $\alpha_{31} = -0.047 \eta^3$, we obtain from the first two of eqs. (45) $\lambda_1 = 2.9 (M/\mu)$, $\lambda_2 = 1.3 M/\mu$. With these values of the parameters, we find a value of α_{11} which is quite close to zero at low energies.

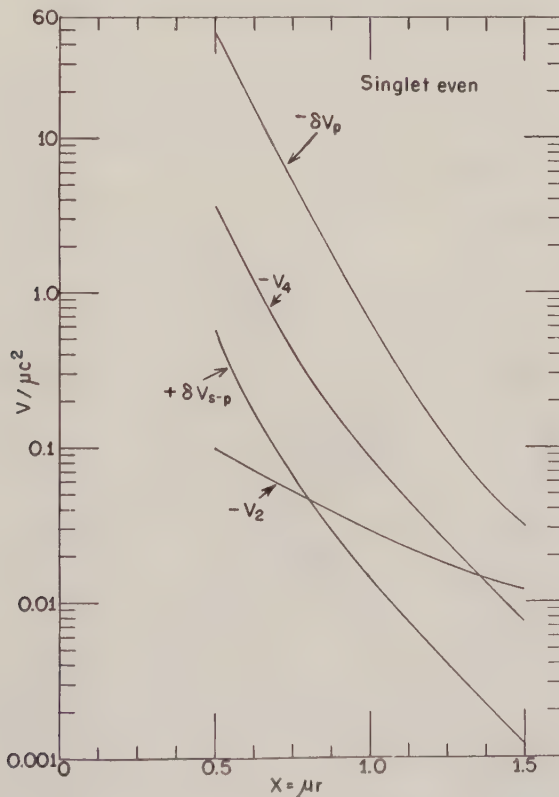


Fig. 2. Central force in singlet even states. Shown separately are the second order potential (V_2), fourth order perturbation result (V_4), P -wave scattering corrections (V_P) and S - P interference terms (V_{S-P}). The total potential which is the algebraic sum of these contributions is not shown.

We now present the fourth order nuclear forces obtained from the combined use of (42), (44), and (45), omitting all details of calculation. We first display the *P*-wave contribution, the origin of the various terms being evident from the coefficients involved, and the notation standard :

$$\begin{aligned} v_{4P}(r) = & (f^2/4\pi)^2/\mu \left\{ \boldsymbol{\tau}^{(1)} \cdot \boldsymbol{\tau}^{(2)} \left(-\frac{8}{\pi} \right) \left[K_1(2x) \left(\frac{23}{4x^4} + \frac{3}{x^2} \right) + K_0(2x) \left(\frac{23}{4x} + \frac{1}{x} \right) \right] \right. \\ & + \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} \left(\frac{8}{\pi} \right) \left[K_1(2x) \left(\frac{3}{x^4} + \frac{2}{x^2} \right) + K_0(2x) \frac{3}{x^3} \right] \\ & + S_{12} \left(-\frac{8}{\pi} \right) \left[K_1(2x) \left(\frac{15}{4x^4} + \frac{1}{x^2} \right) + K_0(2x) \frac{3}{x^3} \right] \Big\} \\ & - (f^2/4\pi)^2/\mu \frac{1}{x^3} \left(\frac{2}{\pi} \right) \left\{ (3 - 2\boldsymbol{\tau}^{(1)} \cdot \boldsymbol{\tau}^{(2)}) \left[\frac{(4 - 4x - x^2)}{x} K_1(x) e^{-x} \right. \right. \\ & \left. \left. + (2 + 2x + x^2) K_0(x) e^{-x} \right] + \frac{2}{3} \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} (3 - 2\boldsymbol{r}^{(1)} \cdot \boldsymbol{r}^{(2)}) \times \right. \end{aligned}$$

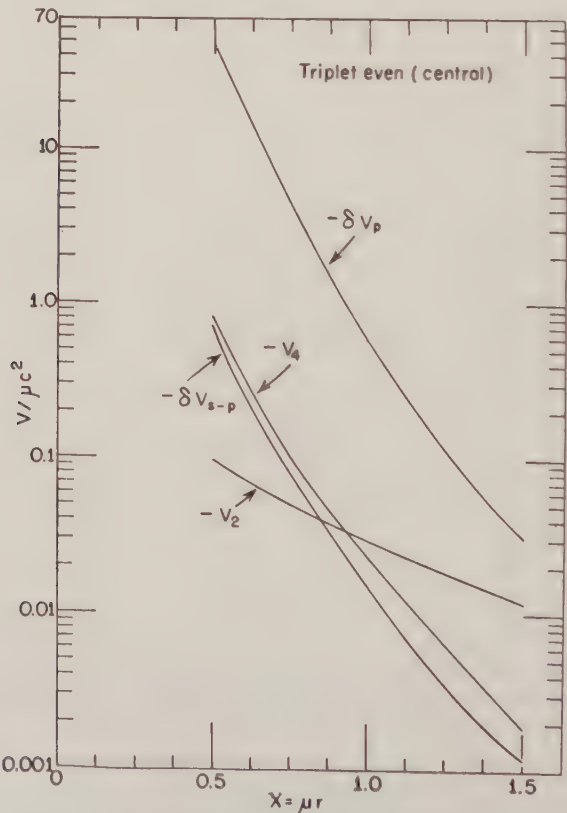


Fig. 3. Central force in triplet even states. Contributions have the same relative significance as in Fig. 2.

$$\begin{aligned}
 & \times \left[\frac{2+2x+x^2}{x} K_1(x) e^{-x} + (1+x) K_0(x) e^{-x} \right] \\
 & - \frac{1}{3} S_{12} (3 - 2 \boldsymbol{\tau}^{(1)} \cdot \boldsymbol{\tau}^{(2)}) \left[\frac{(5+5x+x^2)}{x} K_1(x) e^{-x} + (1+x) K_0(x) e^{-x} \right] \Big\} \\
 & - 3 \lambda_1^2 (f^2/4\pi)^2 (\mu/2M)^2 \mu \left(\frac{2}{\pi} \right) \\
 & \times \left\{ K_1(2x) \left(\frac{43}{2x^5} + \frac{63}{4x^4} + \frac{1}{x^2} \right) + K_0(2x) \left(\frac{43}{2x^5} + \frac{5}{x^3} \right) \right\} \\
 & + \lambda_2^2 (f^2/4\pi)^2 (\mu/2M)^2 \mu \left(\frac{2}{\pi} \right) \boldsymbol{\tau}^{(1)} \cdot \boldsymbol{\tau}^{(2)} \\
 & \times \left\{ \frac{4}{3} \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} K_1(2x) \left(\frac{10}{x^5} + \frac{7}{x^4} \right) + K_0(2x) \left(\frac{10}{x^5} + \frac{2}{x^3} \right) \right\} \\
 & - \frac{2}{3} S_{12} \left[K_1(2x) \left(\frac{35}{2x^5} + \frac{10}{x^4} \right) + K_0(2x) \left(\frac{35}{2x^5} + \frac{2}{x^3} \right) \right]
 \end{aligned}$$

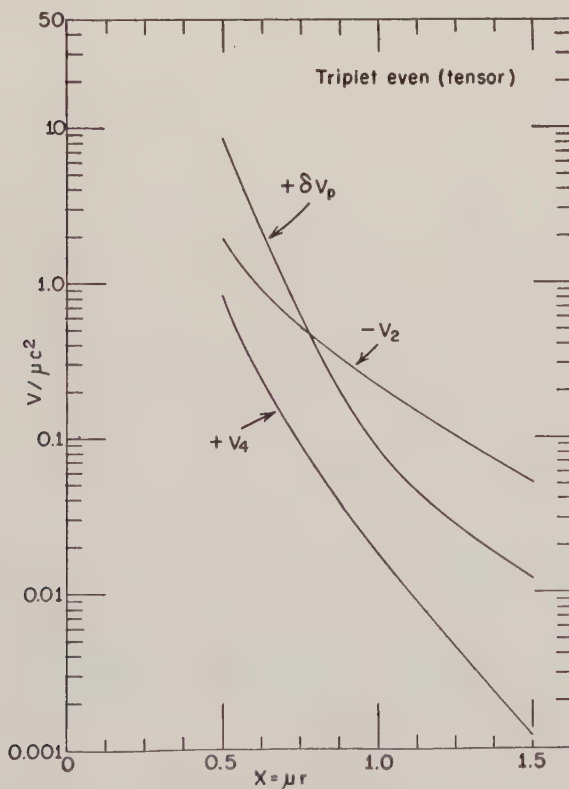


Fig. 4. Tensor force in triplet even states. Notation same as in Fig. 3, except that there is no S - P contribution.

$$\begin{aligned}
& -6\lambda_1 \left(\frac{f^2}{4\pi} \right)^2 (\mu/2M) \mu (e^{-2x}/x^2) \left[\frac{6}{x^4} + \frac{12}{x^3} + \frac{10}{x^2} + \frac{4}{x} + 1 \right] \\
& + \lambda_2 (f^2/4\pi)^2 (\mu/2M) \mu \left\{ \frac{8}{3} \sigma^{(1)} \cdot \sigma^{(2)} \tau^{(1)} \cdot \tau^{(2)} (e^{-2x}/x^2) \left[\frac{3}{x^4} + \frac{6}{x^3} + \frac{5}{x^2} + \frac{2}{x} \right] \right. \\
& \left. - \frac{4}{3} \tau^{(1)} \cdot \tau^{(2)} S_{12} (e^{-2x}/x^2) \left[\frac{6}{x^4} + \frac{12}{x^3} + \frac{8}{x^2} + \frac{2}{x} \right] \right\}. \quad (46)
\end{aligned}$$

For the S - P interference contribution, we find

$$\begin{aligned}
v_{4S-P}(r) &= 6(f^2/4\pi)(g^2/4\pi)\rho_1(\mu/2M)\mu(1+x^{-1})^2(e^{-2x}/x^2) \\
& + (f^2/4\pi)^2\rho_1\rho_2(\tau^{(1)} \cdot \tau^{(2)}) \left(\frac{8}{\pi} \right) \left[K_1(2x) \left(\frac{5}{2x^4} - \frac{1}{x^2} \right) - K_0(2x) \frac{5}{2x^3} \right] \\
& + (f^2/4\pi)^2\lambda_1\rho_1\mu \left(\frac{12}{\pi} \right) \left[K_1(2x) \left(\frac{5}{2x^4} + \frac{1}{x^2} \right) + K_0(2x) \frac{5}{2x^3} \right]. \quad (47)
\end{aligned}$$

We notice that λ_2 does not occur in these terms.

Finally, for the pure S -wave contribution, we obtain

$$\begin{aligned}
v_{4S}(r) &= -3(g^2/4\pi)^2(\mu/2M)^2\mu\rho_1^2 \left(\frac{2}{\pi} \right) K_1(2x)/x^2 \\
& + (g^2/4\pi)^2(\mu/2M)^4\mu\rho_1^2 \left(\frac{2}{\pi} \right) \tau^{(1)} \cdot \tau^{(2)} \left[\frac{K_1(2x)}{x^4} + \frac{K_0(2x)}{x} \right]. \quad (48)
\end{aligned}$$

Of these results the first term of eq. (48) was derived some time ago¹⁰ by means of a special form of the method of this section and used to prove that the S -wave contribution was indeed negligible compared with that of the P -wave. This continues to be true even after the addition of the second isotopic spin term. Because of the exceedingly small value of ρ_1 and small value of ρ_2 the S - P interference potential is also small compared with the pure P -terms. To illustrate this conclusion and to exhibit the relative importance of the P -terms, especially the "scattering corrections", we have plotted the central force in singlet even states in Fig. 2, exhibiting the various contributions, second order, fourth order perturbation, P -waves and S - P wave "scattering correction" separately. The corresponding situation for the central and tensor force in triplet even states is exhibited in Figs. 3 and 4 respectively. A detailed analysis of the potential in Sec. 2 shows that the effective range approximation overestimates the actual P -wave contribution by at least a factor of two and in some cases by close to an order of magnitude. Nevertheless, Fig. 2 establishes unequivocally the importance of the additional P -wave contributions to the central force. On the other hand Fig. 4 illustrates that the second order tensor force is still of paramount importance.

References

- 1) A. Klein *Prog. Theor. Phys.* **20** (1958), 257.
- 2) A. Klein and B. H. McCormick, *Phys. Rev.* **104** (1956), 1747.
- 3) H. Miyazawa, *Phys. Rev.* **104** (1956), 1741.
- 4) *Proceedings Seventh Annual Rochester Conference*, Sec. IV, Interscience Publishers, New York, (1957).
- 5) Hiida, Iwadare, and Machida, *Prog. Theor. Phys.* **15** (1956), 189.
- 6) A. Klein, *Phys. Rev.* **90** (1953), 1101.
- 7) R. Oehme, *Phys. Rev.* **102** (1956), 1174.
- 8) A. Klein, *Phys. Rev.* **104** (1956), 1131.
- 9) J. Schwinger, *Proc. Natl. Acad. Sci. U. S.* **37** (1951), 452.
- 10) M. Gell-Mann and F. E. Low, *Phys. Rev.* **84** (1951), 350.
- 11) A. Klein, *Phys. Rev.* **95** (1954), 1061.
- 12) A. Klein, *Phys. Rev.* **99** (1955), 998.
- 13) J. Orear, *Nuovo Cimento* **4** (1956), 856.
- 14) G. I. Chew and F. E. Low, *Phys. Rev.* **101** (1956), 1570.
- 15) S. Barnes, *Proceedings Seventh Annual Rochester Conference*, Sec. II.

Magnetic Properties of the Iron-Group Anhydrous Chlorides

Junjiro KANAMORI

Department of Physics, Osaka University, Osaka

(Received August 27, 1958)

The magnetic properties of FeCl_2 are investigated in detail from the standpoint of the one-ion approximation. The lowest orbital state of the ferrous ion in the crystalline field of trigonal symmetry is assumed to be doublet, in which the component of the orbital angular momentum along the trigonal axis is not quenched. The spin-orbit coupling energy produces a splitting of the spin levels belonging to the lowest doublet. In the vicinity of, or below, the Néel temperature (24°K) where only the lowest spin level is populated, we have effectively an Ising model, in which the x, y components of the spins are completely quenched. The metamagnetic behaviour of this substance at liquid hydrogen temperatures can quite naturally be interpreted by this model. Also the susceptibility near the room temperature is discussed and it is shown that the paramagnetic Curie temperature is strongly modified by the large splitting of the spin levels due to the spin-orbit coupling energy. The crystalline field is calculated, and the result seems to support the assumed orbital level scheme. FeCO_3 , CoCl_2 , NiCl_2 , FeBr_2 and FeI_2 are also discussed briefly.

§ 1. Introduction

The anhydrous chlorides of the iron-group elements are sometimes grouped as metamagnetic substances, because their susceptibility is more or less field-dependent at low temperatures.¹⁾ Especially, FeCl_2 is known to have a strongly field-dependent susceptibility at liquid hydrogen temperature.²⁾ The purpose of this paper is to make a theoretical investigation of FeCl_2 , based on one-ion approximation which has hitherto been applied to several antiferromagnetic ionic crystals^{3,4)} with considerable success, and further to discuss the magnetic properties of FeBr_2 , FeI_2 , CoCl_2 and NiCl_2 from the same point of view.

In 1933, Landau⁵⁾ suggested that the field-dependence of the susceptibility of these chlorides might be related to the transition from antiferromagnetic to ferromagnetic state at a certain external field. All these three chlorides have the crystal structure of the CdCl_2 type,⁶⁾ which, if one neglects a small deformation, can be derived from the lattice of the NaCl type by taking away alternate hexagonal close-packed layers of cations. Landau assumed that ferromagnetic interactions align the magnetic moments parallel within each hexagonal layer, and that, in the absence of an external field a comparatively weak antiferromagnetic interaction makes the moments of adjacent layers antiparallel. This picture has recently been confirmed by neutron diffraction experiments,⁷⁾ which also have shown that the moments are oriented along the trigonal axis in FeCl_2 but they are perpendicular to this axis in CoCl_2 .

According to magnetic measurements on powder specimens by Starr, Bitter and

Kaufmann,³⁾ the susceptibilities of FeCl_2 , CoCl_2 and NiCl_2 under low field show a maximum, which is characteristic of antiferromagnetism, at 24°K, 25°K and 50°K respectively, where also specific heat anomalies are observed.³⁾ The magnetization curve of FeCl_2 below 24°K is characterized by a small initial susceptibility which decreases rapidly with decreasing temperature, followed by a very large differential susceptibility and eventual approach to magnetic saturation. The saturation of magnetization at high field suggests the occurrence of ferromagnetism. On the other hand, in CoCl_2 and NiCl_2 the susceptibility shows only a small dependence both on temperature and on field strength below the Néel temperature. Recently, Bizette et al.¹⁰⁾ have made magnetic measurements on single crystals of these substances, whose results will be referred to in later sections (§ 4 and § 7).

Néel¹⁰⁾ has given a phenomenological explanation of these properties. He ascribed the behaviour of FeCl_2 to the transition from antiferromagnetic to ferromagnetic state, whereas for the cases of CoCl_2 and NiCl_2 he assumed that the magnetic moments flip from the easy axis to the direction perpendicular to the applied field in the layers. From energy consideration he concluded that in FeCl_2 a strong anisotropy energy binding the moments to the easy axis must be present.

In this paper we shall show, by investigating the microscopic origin of the anisotropy energy, that the experimental data on FeCl_2 both in antiferromagnetic and paramagnetic states can be reasonably accounted for. We show also that the difference between the magnetic properties of FeCl_2 and those of CoCl_2 can be explained by such an investigation. Our main idea is as follows:—

Since the chlorine ions form approximately a face-centred cubic lattice, the main part of the crystalline field around a cation is of cubic symmetry. In addition to it, there is a crystalline field of trigonal symmetry which arises from the non-cubic arrangement of cations, the induced electric dipole moments of the chlorine ions and the small departure from cubic symmetry in the chlorine ion arrangement. In the cubic field, the ground orbital state of a free ferrous ion, 5D , splits into two sublevels, of which the lower one, Γ_5 , is triply degenerate.¹¹⁾ The trigonal field splits this sublevel into a doublet and a singlet (see Fig. 1). The main origin of the magnetic anisotropy energy is considered to be the spin-orbit coupling modified by the crystalline field.¹²⁾ We assume that the lowest orbital state is the doublet, because the anisotropy energy favouring the trigonal axis cannot be obtained with the singlet as the lowest state. (This is reversed in CoCl_2). Further we assume that the doublet is well separated in energy from the singlet as compared with the spin-orbit coupling energy.*

With these assumptions, we can show that a large splitting is produced by the spin-orbit coupling energy among the spin levels belonging to the orbital doublet, since the component of the orbital angular momentum along the trigonal axis is not completely quenched. The lowest spin level is doublet, in which the z -component of the spin takes

* As will be discussed in later sections, we cannot judge from the experimental results obtained hitherto whether this assumption is valid or not. However, the calculations in the following sections will serve at least to the qualitative understanding, even when future experiments invalidates this assumption. The calculations without this assumption will be published in another paper.

values ± 2 , while the x and y components are completely quenched. (Here we take the trigonal axis as the z -axis.) In the vicinity of, and below the Néel temperature, where only the lowest spin doublet is populated, the Ising model is therefore realized for the spin system of FeCl_2 , the magnetic moments being tightly bound to the trigonal axis. In the following sections, we discuss the behaviour of the principal susceptibilities both in antiferromagnetic and paramagnetic states and the magnetization curve in antiferromagnetic state. The theory can well explain the data of the powder specimen of FeCl_2 , though there remains some ambiguous points in the quantitative comparison with experiments. Also the theory can explain qualitatively the behaviour of the principal susceptibilities of FeCO at room temperature.

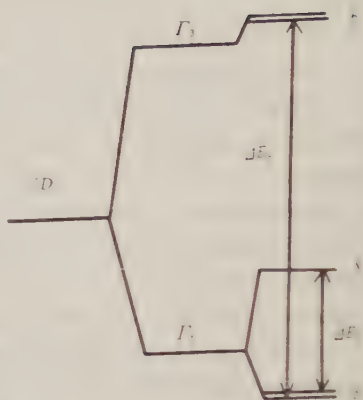


Fig. 1. Orbital level scheme of Fe^{3+} .

An attempt was made to calculate the trigonal field, in order to justify the orbital level scheme assumed above. The result, however, was found to depend sensitively on the magnitude of the effective point dipole moments induced on chlorine ions. If we assume the magnitude of it calculated with the use of the polarizability and the magnitude of the electric field acting on each chlorine ion, the energy of the orbital singlet is found to be lower than that of the doublet. With a magnitude of the induced dipole moment less than about two-third of the calculated value, the assumed orbital level of Fe^{3+} is obtained.

§ 2—§ 6 will deal with FeCl_2 . In § 2, we derive the effective Hamiltonian of the spin-orbit coupling energy and the Zeeman energy for the ground orbital doublet. Using the results of § 2, we discuss the antiferromagnetic state and the transition from the antiferromagnetic to ferromagnetic state in § 3 and § 4; the paramagnetic state is discussed in § 5 and § 6. In the final section we discuss the properties of CoCl_2 , NiCl_2 , FeBr_2 and FeI_2 . Appendices supplement the calculation of the text and give the calculation of the crystalline field.

§ 2. Effective Hamiltonian for the lowest orbital doublet of the ferrous ion

With the assumptions mentioned in § 1, we shall derive the effective Hamiltonian of the spin-orbit coupling energy and Zeeman energy for the orbital doublet of the ferrous ion by the perturbation method, in a similar way to Pryce's^{1,2)} procedure in non-degenerate case.

We take the trigonal axis as the axis of quantization and write φ_M the orbital wave function specified by the azimuthal quantum number M of the orbital angular momentum. The orbital wave functions of the lowest doublet can be expressed as

$$\begin{aligned}\psi'_{g1} &= \cos \vartheta \varphi_2 + \sin \vartheta \varphi_{-1}, \\ \psi'_{g2} &= \cos \vartheta \varphi_{-2} - \sin \vartheta \varphi_1.\end{aligned}\quad (1)$$

The angle ϑ depends on the ratio between the magnitude of the cubic field and that of the trigonal field. If we neglect the effect of the trigonal field, ϑ is $\tan^{-1} 1/\sqrt{2}$. The wave function of the singlet is given by

$$\psi_s = \varphi_0. \quad (2)$$

The wave functions of the highest orbital doublet are given by

$$\begin{aligned}\psi'_{h1} &= \sin \vartheta \varphi_2 - \cos \vartheta \varphi_{-1}, \\ \psi'_{h2} &= \sin \vartheta \varphi_{-2} + \cos \vartheta \varphi_1.\end{aligned}\quad (3)$$

With these ψ' , we can show that in the ground doublet L_z is diagonal and $(g1|L_z|g1) = -(g2|L_z|g2)$, and other components of \mathbf{L} have no matrix elements. The details are given in App. I.

It is convenient to introduce the Pauli σ_z -matrix which operates in the two-dimensional space spanned by ψ'_{g1} and ψ'_{g2} . The first order spin-orbit coupling energy in the ground doublet is expressed as

$$\lambda \langle L \rangle \sigma_z S_z \quad (4)$$

where $\langle L \rangle = (g1|L_z|g1)$. $\langle L \rangle$ is equal to one, if we assume $\vartheta = \tan^{-1} 1/\sqrt{2}$ (see Appendix I). The first order term of the Zeeman energy is given by

$$\mu_R (2S_z + \langle L \rangle \sigma_z) H_z + 2\mu_R (H_x S_x + H_y S_y). \quad (5)$$

The first order energy of the spin-orbit coupling, (4), splits the ground orbital doublet into five doublets, specified by

$$\sigma_z = 1, S_z = m \quad \text{and} \quad \sigma_z = -1, S_z = -m \quad (m=2, \dots, -2).$$

They have a constant spacing $|\lambda| \langle L \rangle$, and the lowest doublet corresponds to $\sigma_z = 1, S_z = 2$ and $\sigma_z = -1, S_z = -2$.

The non-diagonal matrix elements of \mathbf{L} connecting the lowest orbital doublet to higher states produce the second order perturbation energies of the spin-orbit and orbital Zeeman energies, which contain non-diagonal as well as diagonal terms with respect to the spin doublets produced by the first order energy, (4). Since the non-diagonal terms are not important in the following discussions because they are small, we pick up here only the diagonal terms. They are given by

$$\begin{aligned}H_1 &= -(3\lambda^2 \sin^2 \vartheta / 2 \Delta E_s) (S_x^2 + S_y^2 + \sigma_z S_z) \\ &\quad - (\lambda^2 / \Delta E_h) \{ (9/4) \sin^2 2\vartheta S_x^2 + S_x^2 + S_y^2 - \sigma_z S_z \},\end{aligned}\quad (6)$$

$$\begin{aligned}H_2 &= (3\mu_R |\lambda| \sin^2 \vartheta / \Delta E_s) (H_x S_x + H_y S_y) + (\mu_R |\lambda| / \Delta E_h) \\ &\quad \times \{ (9/2) \sin^2 2\vartheta H_z S_z + 2(H_x S_x + H_y S_y) \},\end{aligned}\quad (7)$$

and

$$H_3 = - (3\mu_B^2 \sin^2 \vartheta / 2 \Delta E_s) (H_x^2 + H_y^2) - (\mu_B^2 / \Delta E_h) \times \{ (9/4) \sin^2 2\vartheta H_z^2 + H_x^2 + H_y^2 \}, \quad (8)$$

where ΔE_s is the energy separation between the lowest orbital doublet and the orbital singlet and ΔE_h the separation between the two orbital doublets. H_1 is the second order energy of the spin-orbit coupling, H_2 the cross term of the orbital Zeeman energy and spin-orbit coupling energy, and H_3 the second order term of the orbital Zeeman energy. H_2 gives the $(g-2)$ part of the Zeeman energy, while H_3 gives rise to the temperature independent susceptibility.

§ 3. Theory of the antiferromagnetic state

If we assume $\langle L \rangle = 1$ and the free ion value of λ , -100 cm^{-1} , the spacing of the doublets will be $|\lambda| \langle L \rangle = 100 \text{ cm}^{-1}$, or 144°K . If ΔE_s is larger than 1000 cm^{-1} , the second order term, H_1 , affects this separation only by an amount less than 10 per cent. Therefore, in the vicinity of the Néel temperature, 24 K . or below it, almost only the lowest doublet is populated. In this lowest doublet, S_+ and S_- have no matrix element, and S_z and σ_z can be replaced by the effective spin operator s , of magnitude $1/2$ through the relations,

$$S_z = 4s_z \quad \text{and} \quad \sigma_z = 2s_z. \quad (9)$$

Thus we retain only the diagonal part of the Hamiltonian with respect to the lowest spin doublet, and such a truncated Hamiltonian is essentially the Hamiltonian of the Ising model.

The interaction between the magnetic moments consists of the exchange interaction and the magnetic dipole-dipole interaction. The latter is small, but not negligible compared with the former. For the exchange interaction we take the Hamiltonian

$$H_{ex} = -2J_1 \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + 2J_2 \sum_{\langle k,l \rangle} \mathbf{S}_k \cdot \mathbf{S}_l, \quad (10)$$

which reduces for the lowest doublet to

$$H_{ex} = -32J_1 \sum_{\langle i,j \rangle} s_{iz} \cdot s_{jz} + 32J_2 \sum_{\langle k,l \rangle} s_{kz} \cdot s_{lz}.$$

The first sum is the ferromagnetic interaction within the same hexagonal layer, the second sum represents the antiferromagnetic inter-layer interactions, and $\langle \cdot, \cdot \rangle$ means a pair. In the following, we shall treat these interactions in the molecular field approximation.

The effective field arising from the magnetic dipole-dipole interaction can be divided into three parts: the first is the field arising from the magnetic moments within the Lorentz sphere, the second the Lorentz field, and the third the demagnetizing field. The total effective dipole field acting on each ion belonging to the + sublattice is therefore

$$H_d^+ = \phi_1 M^+ + \phi_2 M^- + (4\pi/3) (M^+ + M^-) - N(M^+ + M^-), \quad (11)$$

where ϕ_1 and ϕ_2 are the diagonal tensors obtained from the dipole sum within the Lorentz sphere, N is the tensor of the demagnetizing factor, and M^+ and M^- are the magnetizations of the sublattices per c. c. The magnitudes of ϕ_1 and ϕ_2 are given in Appendix II. For the ions of the — sublattice ϕ_1 and ϕ_2 should be interchanged.

We discuss first the principal susceptibilities in low field. As discussed before, the x , y components of the applied field disappear in the Hamiltonian of the first approximation. However, the non-diagonal elements of S_x and S_y between the lowest doublet and the second doublet of $S_z=1$, $\sigma_z=1$ and $S_z=-1$, $\sigma_z=-1$ give rise to the second order perturbation energy, from which we obtain

$$\chi_{\perp} = 2Ng_x^2 \mu_B^2 / |\lambda| \langle L \rangle. \quad (12)$$

Here we neglected the contribution arising from H_3 given by (8), as it turns out to be small.

For the susceptibility along the trigonal axis we expect the normal behaviour of the antiferromagnetic parallel susceptibility. In the molecular field approximation, the effective fields to which the effective spins of the \pm sublattices are subject are given by

$$H_{eff}^{\pm} = H + (\Gamma + \phi_{1z} + 4\pi/3 - N_z)M^{\pm} - (A - \phi_{2z} - 4\pi/3 + N_z)M^{\mp}, \quad (13)$$

where A and Γ are defined by

$$\Gamma = 64J_1z_1/N_0(4g_z + 2\langle L \rangle)^2 \quad \text{and} \quad A = 64J_2z_2/N_0(4g_z + 2\langle L \rangle)^2.$$

z_1 and z_2 are the numbers of the neighbours which interact with each ion and N_0 is the number of cations per c. c.

The parallel susceptibility at the Néel temperature is given by

$$\chi_{\parallel mol} |_{T=T_N} = (N/N_0) \cdot (A - \phi_{2z} - 4\pi/3 + N_z)^{-1}. \quad (14)$$

Below the Néel temperature we expect that the parallel susceptibility tends to zero with decreasing temperature as in ordinary antiferromagnetic substances.

At a high field, a transition from antiferromagnetic to ferromagnetic state takes place. This transition was discussed previously by the present author, Motizuki and Yosida⁽¹⁰⁾ and also by Gorter and van Peski-Tinbergen,⁽¹¹⁾ both using the molecular field approximation. If the coefficients of the molecular field satisfy the relation,

$$1 \geq (\Gamma + \phi_{1z} + 4\pi/3 - N_z) / (A + \Gamma + \phi_{1z} - \phi_{2z}) \geq 3/8, \quad (15)$$

the first order transition takes place in the temperature range given by

$$0 \leq T/T_N \leq 4/3 - (1/3) (A + \Gamma + \phi_{1z} - \phi_{2z}) / (\Gamma + \phi_{1z} + 4\pi/3 - N_z). \quad (16)$$

At absolute zero, the transition takes place when the strength of the field applied along the trigonal axis reaches a value,

$$H_{c0} = (A - \phi_{2z} - 4\pi/3 + N_z)M, \quad (17)$$

where $M=|M^\perp|$. The transition field decreases with increasing temperature.*

If the applied field makes an angle of ϑ with the trigonal axis, the transition field is given by

$$H_c = H_{c0}/\cos\vartheta, \quad (18)$$

because in this case x and y components of the field have no effect. However, for a powder specimen, we must be careful of the demagnetizing field. If we assume simply that each microcrystal is subject to an average demagnetizing field, which is determined by the total magnetization of the specimen, and an appropriate demagnetizing factor N_p , the transition field for those microcrystals, whose trigonal axis makes an angle of ϑ with the applied field, will be given by

$$H_c = N_p I(H_c) + H_{c0}'/\cos\vartheta, \quad (19)$$

where $I(H_c)$ is the average magnetization per c. c. of the powder specimen at the field H_c and H_{c0}' is the transition field we would have for $\vartheta=0$ if there were no demagnetizing field. The microcrystals with an angle of ϑ contribute to $I(H)$ a quantity $2M \cos\vartheta$ after the transition. So, if we denote the probability of their occurrence by $n(\vartheta)$, $I(H)$ will be given by

$$I(H) = 2M \int_{H_{c0}'}^H n(\vartheta) \cos\vartheta (d\vartheta/dH_c) dH_c, \quad (20)$$

neglecting the magnetization arising from the antiferromagnetic state. If we assume $n(\vartheta) = \sin\vartheta$, i. e., the random distribution of the easy axis of the microcrystals, (19) and (20) give the relation,

$$I(H) = M \{1 - H_{c0}'^2/(H - N_p I(H))^2\}. \quad H \geq H_{c0}' \quad (21)$$

From (21) we see that the magnetization of the powder specimen is zero at $H = H_{c0}'$ (actually small), begins to increase linearly with increasing field beyond H_{c0}' , and then tends to saturate to M . In a single crystal this takes place at a higher field given by $H_{c0} = H_{c0}' + N_z M$. This conclusion is, however, not decisive since the assumption of an average demagnetizing field might not be justified in the initial stage of the magnetization.

§ 4. Comparison of the theory of the preceding section with experiments

First we discuss the principal susceptibilities.

In order to estimate the order of magnitude of χ_\perp , we assume $|\lambda| = 100 \text{ cm}^{-1}$ and $g_e = 2.1$, the latter corresponding to $\Delta E_s = 1,000 \text{ cm}^{-1}$. Then (12) gives

$$\chi_{\perp \text{ mol}} = 0.023. \quad (22)$$

* Yomosa and Suzuki have discussed this temperature effect on the transition field, using the Bethe approximation as well as the molecular field approximation. (S. Yomosa and H. Suzuki, yet unpublished.)

Further, if we assume $\langle L \rangle = 1$ and $g_z = 2.0$, the sublattice magnetization will be

$$M = (N_0 \mu_B / 2) (2g_z + \langle L \rangle) = 357 \phi \text{ per c. c.} \quad (23)$$

$$= 13,970 \phi \text{ per mol.}$$

If we assumed $\langle L \rangle = 1.5$, we would have obtained $M = 15,400 \phi$ per mol. The $\langle L \rangle$ value and g -value will be discussed in later sections. Here we use (23) preliminarily for order estimation.

Fig. 2 shows the experimental curve²⁾ of magnetization vs. field at 13.9°K, from which we estimate the transition field to be about 10,000 ϕ . Using (23) and (14), we have

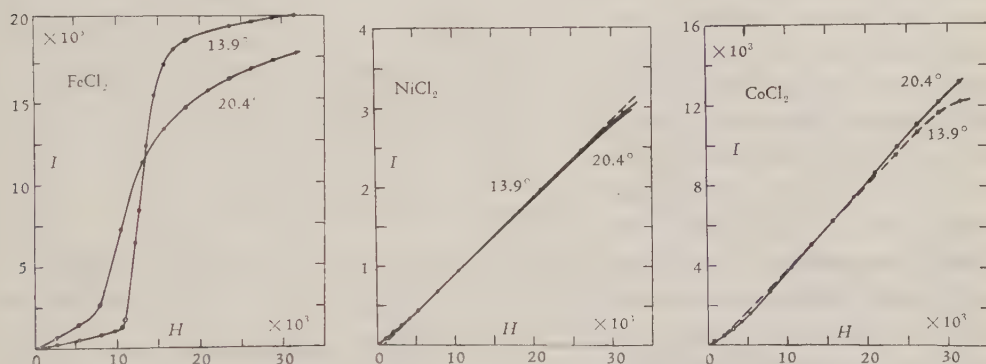


Fig. 2. Experimental curves of magnetization vs. field of FeCl_2 , CoCl_2 and NiCl_2 below their Néel temperatures.²⁾

$$H_{cr}' / M = A - \Phi_{2z} - (4\pi/3) = 10,000 / 357 = 28.0 \quad (24)$$

and

$$\chi_{\parallel mol} |_{T_N} = 1.4. \quad (25)$$

In (25), we assumed $N_z = 0$. If we assume $N_z = 4\pi$, we would obtain $\chi_{\parallel mol} |_{T_N} = 0.97$. From (24) and $\Phi_{2z} = 8.37$ (see (A-3)) we have $A = 40.6$, so that in FeCl_2 the exchange interaction between the layers is larger than the magnetic dipole-dipole interaction between the layers.

From (22) and (25) we obtain for the powder susceptibility

$$\chi_{p mol} |_{T_N} = 0.48.$$

The experimental value of $\chi_{p mol} |_{T_N}^{2)}$ is about 0.4. In the vicinity of the Néel temperature, the theory predicts that χ_{\parallel} is much greater than χ_{\perp} and therefore the powder susceptibility is approximately given by $\chi_{\parallel}/3$, which is in accordance with the experimental result that χ_p decreases rapidly with decreasing temperature.

With the assumption of the Ising model and molecular field, the Néel temperature is calculated by

$$kT_N = (N_0/2) (2g_z + \langle L \rangle)^2 \mu_B^2 (A + \Gamma + \Phi_{1z} - \Phi_{2z}).$$

Inserting $T_N = 24$ K, we see that the relation (15) is satisfied and that (16) becomes

$$0 \leq T/T_N \leq 0.95.$$

We show in Fig. 3 the theoretical curve of magnetization vs. field at absolute zero assuming $N_p = 0$. This assumption is not very serious, since we see from (23) that the contribution of the demagnetizing field is at most $4\pi M = 4500 \phi$. The general behaviour of this curve is in accordance with the observation by Starr, Bitter and Kaufmann.²⁾ However, the theory predicts about 14,000 ϕ per mol for the saturation value of the magnetization, while the experiment gives about 20,000 ϕ per mol. This discrepancy may be due to the following. The experimental value was obtained after repeated magnetizations in a high field in order to eliminate a hysteresis observed in the initial stage of the experiment. Since this hysteresis was due to change of orientation of the microcrystals on magnetization, the assumption of the random orientation of microcrystals is inappropriate. Also, the assumed values of $\langle L \rangle$ and g might be somewhat in errors.

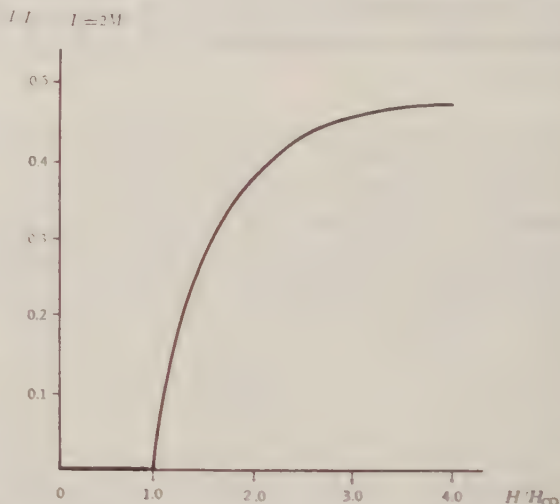


Fig. 3. Theoretical curve of magnetization vs. field of FeCl_2 at 0°K .

Recently Bizette et al.¹⁰⁾ made magnetic measurements on a single crystal specimen. They found that the magnetization along the trigonal axis attains to a value of 36,500 ϕ per mol under a field of 24,000 ϕ at 4 K and does not yet saturate. This is quite unintelligible theoretically, since even if we assumed the maximum possible value of 2 for $\langle L \rangle$, the saturation magnetization would have been no greater than 33,500 ϕ per mol. The present theory seems also incompatible with their experiment in that the observed perpendicular susceptibility is larger by one order compared with (23).

§ 5. Theory of the paramagnetic state

We discuss the paramagnetic state on the basis of the following simplified one-ion Hamiltonian,

$$H = \lambda \langle L \rangle S_z \sigma_z - 2Jz \langle S \rangle \cdot S + \mu_B H_z (g_z S_z + \langle L \rangle \sigma_z) + \mu_B g_x (H_x S_x + H_y S_y), \quad (26)$$

where $2Jz$ is the coefficient of the molecular field effective in the paramagnetic state. We

neglected the dipole interaction in (26) because their effect is small. $\langle S \rangle$ is the average value of the spin which is to be determined self-consistently. We define m and m' as quantum numbers of S_z and σ_z . Then the susceptibility along the trigonal axis is given by

$$\chi_{||} = (N\mu_B^2/kT) \{ \langle L \rangle^2 + 2g_z \langle L \rangle \langle mm' \rangle + g_z^2 \langle m^2 \rangle + \langle L \rangle^2 (2J_z/kT) (\langle mm' \rangle^2 - \langle m^2 \rangle) \} / \{ 1 - (2J_z/kT) \langle m^2 \rangle \}, \quad (27)$$

where $\langle \rangle$ means the thermal average. $\langle m^2 \rangle$ and $\langle mm' \rangle$ are given by

$$\langle m^2 \rangle = 2 \sinh(x/2) (4 \cosh 2x + \cosh x) / \sinh(5x/2), \quad (28)$$

$$\langle mm' \rangle = 2 \sinh(x/2) (2 \sinh 2x + \sinh x) / \sinh(5x/2), \quad (29)$$

with $x = |\lambda| \langle L \rangle / kT$.

In a similar way, the perpendicular susceptibility is obtained as

$$\chi_{\perp} = (2Ng_z^2 \mu_B^2/kT) f(x) / \{ 1 - (4J_z/kT) f(x) \} \quad (30)$$

with

$$f(x) = \sinh x \sinh(x/2) (4 \cosh x + 1) / x \sinh(5x/2). \quad (31)$$

Since the population in each spin level changes with temperature, there is a specific heat anomaly of the Schottky type, which is given by

$$\Delta C = R \{ (x/2)^2 \operatorname{cosech}^2(x/2) - (5x/2)^2 \operatorname{cosech}^2(5x/2) \}. \quad (32)$$

§ 6. Comparison of the theory of the preceding section with experiment

Neglecting the dipole interaction, we obtain the relations

$$2J_z = 2J_1 z_1 - 2J_2 z_2 \quad (33)$$

and

$$8J_2 z_2 / (2g_z + \langle L \rangle) \mu_B = AM = 10,000 \phi. \quad (34)$$

The Néel temperature is defined as the temperature at which an infinitesimal antiferromagnetic molecular field sets in. This condition gives

$$kT_N = (2J_1 z_1 + 2J_2 z_2) \langle m^2 \rangle_{T=T_N}. \quad (35)$$

Inserting $T_N = 24^\circ \text{K}$ and assuming $|\lambda| \langle L \rangle = 100 \text{ cm}^{-1}$, we obtain from (28) and (35)

$$2J_1 z_1 + 2J_2 z_2 = 24 k / 3.996 = 8.30 \times 10^{-16} \text{ erg}.$$

It might be remarked that for the perfect Ising model we obtain $2J_1 z_1 + 2J_2 z_2 = 24 k/4$ and for the Heisenberg model $2J_1 z_1 + 2J_2 z_2 = 24 k/2$. Thus in FeCl_2 the large anisotropy energy enhances the Néel temperature twice as high as the value expected without it. From (33), (34) and (35) $2J_z$ is estimated as

$$2Jz = 6.0 \times 10^{-16} \text{ erg}$$

$$\text{or } 3.0 \text{ cm}^{-1}.*$$

It is known experimentally that the powder susceptibility at room temperature follows the Curie-Weiss law²⁾

$$\chi_{p \text{ mol}} = 3.59 / (T - 48), \quad (36)$$

or according to older data obtained by Honda and Ishiware¹⁵⁾

$$\chi_{p \text{ mol}} = 3.51 / (T - 17). \quad (37)$$

Further, according to the data obtained at Leiden,¹⁶⁾ the paramagnetic Curie temperature is equal to 20.4° . It is noted that the value of 48° for θ appearing in (36) is twice as large as the Néel temperature. In the ordinary molecular field approximation we expect that $\theta < T_N$, because, as seen from (33) and (35), the coefficient of the molecular field which is effective in the paramagnetic state is smaller than that which is effective in the antiferromagnetic state. Moreover, we have seen that in this case T_N is enhanced by virtue of the Ising model to twice as high as that expected for the Heisenberg model. Thus in the limit of $kT \gg |\lambda|$, we expect $\theta < T_N/2$. Actually, however, the numerical calculation of (27) and (30) shows that an apparent high paramagnetic Curie

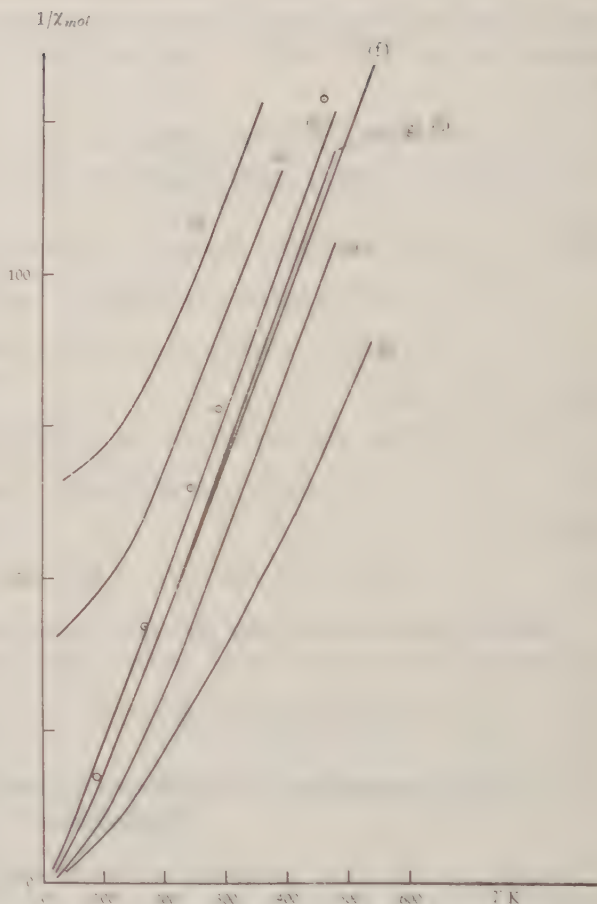


Fig. 4. Susceptibility of FeCl_2 in paramagnetic region.

- curve (a): χ_{\parallel} calculated with $\langle L \rangle = 1$ and $g_z = 2$;
- curve (b): χ_{\perp} calculated with $\langle L \rangle = 1$ and $g_z = 2.1$;
- curve (c): χ_{\parallel}' calculated with $\langle L \rangle = 1.5$ and $g_z = 2.025$;
- curve (d): χ_{\perp}' calculated with $\langle L \rangle = 1.5$ and $g_z = 2.04$;
- curve (e): $(1/3)(\chi_{\parallel} + 2\chi_{\perp})$;
- curve (f): $(1/3)(\chi_{\parallel}' + 2\chi_{\perp}')$;
- curve (g): experimental curve after Starr, Bitter and Kaufmann;
- curve (h): $(1/2)(\chi_{\parallel} + \chi_{\perp})$ (it coincides with (g));
- (•) is the experimental values after Honda and Ishiware.

* One might question whether this procedure of determining $2Jz$ is consistent with the theory of § 3 where the dipole interaction was taken into account. However, it can be shown that, even if we include the magnetic dipole interaction in the calculation, the correction is less than 3 per cent.

temperature results for the susceptibility near room temperature mainly due to the spin-orbit coupling energy, (6).*

If we assume, for instance, $\lambda\langle L\rangle=100\text{ cm}^{-1}$ and $g_z=2.0$, we obtain for χ_{\parallel} near room temperature an approximate formula

$$\chi_{\parallel}=3.85/(T-85). \quad (38)$$

Further neglecting the exchange interaction, we have

$$\chi_{\parallel}=3.85/(T-80). \quad (39)$$

For χ_{\perp} , we obtain, assuming $g_z=2.1$, $\Delta E_s=1,000\text{ cm}^{-1}$ and $\Delta E_h=8,000\text{ cm}^{-1}$,

$$\chi_{\perp}=3.88/(T+64.9). \quad (40)$$

The small temperature-independent susceptibility contributed from H_3 , (8), has been taken into account in (38) (39) and (40).

The principal susceptibilities and the powder susceptibility for various values of the parameters appearing above have been calculated and shown in Fig. 4, together with the experimental powder susceptibility. From this calculation we see that, if we assume $\langle L\rangle=1.5$, a satisfactory agreement is found with the observed susceptibility given by (36), whereas with $\langle L\rangle=1$, we see no good agreement. This conclusion is to some extent indifferent to the molecular field approximation, since the results are insensitive to the choice of the value of $2J_z$. However, the following experimental condition must be noted.

In reference 2), it is stated that the flake shaped microcrystals are not randomly orientated in the vessel but partially aligned putting their trigonal axis parallel to the direction of the applied field. Therefore, it is possible that the measured powder susceptibility is given, for instance, by $(1/2)(\chi_{\parallel}+\chi_{\perp})$ instead of $(1/3)(\chi_{\parallel}+2\chi_{\perp})$. If this were the case, the theoretical susceptibility with $\langle L\rangle=1$ and $g_z=2$ would give a good agreement with (36). On the other hand, the data by Honda and Ishiware, which are summarized by

(37), are rather well represented by the curve calculated with $\langle L\rangle=1$ and $g_z=2$. Although the assumption of $\langle L\rangle=1$ thus appears plausible, we cannot at present definitely

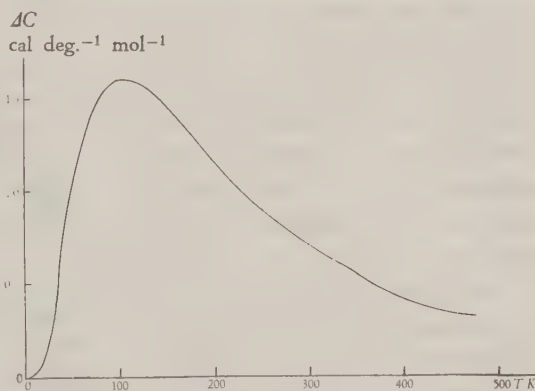


Fig. 5. The anomalous specific heat of FeCl_2 in paramagnetic region (calculated with the assumption of $|\lambda|=100\text{ cm}^{-1}$ and $\langle L\rangle=1$).

* If χ is expanded in powers of $1/T$, the coefficient of the $1/T^2$ term is given by $C\theta$. Thus at high temperatures one would obtain θ by expanding (27) and (30) in powers of $1/T$. We have in this way $\theta=8|\lambda|/9k=128^\circ\text{K}$ ($|\lambda|$ is assumed to be 100 cm^{-1}) with $\langle L\rangle=1$ and $g_z=2$ for χ_{\parallel} and $\theta=0$ for χ_{\perp} . However, this procedure is not justified at room temperature as the series converge poorly.

determine from experimental results the $\langle L \rangle$ value and other parameter values.

The calculated anomalous specific heat given by (32) is shown in Fig. 5.

A little note will be added here as to FeCO_3 . The ion Fe^{2+} in this crystal is subject to a trigonal crystalline field, and this substance is suspected to be antiferromagnetic below 35°K. Foëx¹⁷⁾ has found $\theta = 60^\circ\text{K}$ for χ and $\theta = -103^\circ\text{K}$ for χ_\perp near room temperature, so that the behaviour of the principal susceptibilities is in qualitative agreement with (39) and (40).

§ 7. Other layer crystals

It has been found by neutron diffraction experiment¹⁸⁾ that FeBr_2 has an antiferromagnetic structure similar to that of FeCl_2 , though its crystal structure is of the CdI_2 type in which the bromine ions form a hexagonal close-packed structure. The orientation of the magnetic moments is along the trigonal axis, suggesting the same orbital level scheme for the ferrous ions as in FeCl_2 . Recently Bizette et al.¹⁹⁾ made magnetic measurements on the single crystal of this substance as well as that of FeI_2 . At their Néel temperatures (11°K for FeBr_2 and 10°K for FeI_2) the susceptibility along the c -axis is much greater than the susceptibility perpendicular to the c -axis, as in FeCl_2 .* No metamagnetic behaviour is found in both substances below the Néel temperature. We expect that the transition field H_c for these substances is higher than that for FeCl_2 , since the antiferromagnetic interaction may be stronger in these substances due to smaller electro-negativities of the bromine and iodine ions.

In CoCl_2 and also in NiCl_2 , the anisotropy energy seems to favour the plane perpendicular to the trigonal axis, though we have so far no neutron diffraction investigation for

NiCl_2 . The anisotropy energy within the plane may be so small as to permit the moments to turn in the plane and become perpendicular to the applied field even when the field is weak. With increasing field strength applied in the plane, the specimen will exhibit the perpendicular susceptibility up to a field strength of $H_c = 2M\chi_\perp$,²⁰⁾ at which the magnetization will reach ferromagnetic saturation (see Fig. 6 (a)). Since the effective

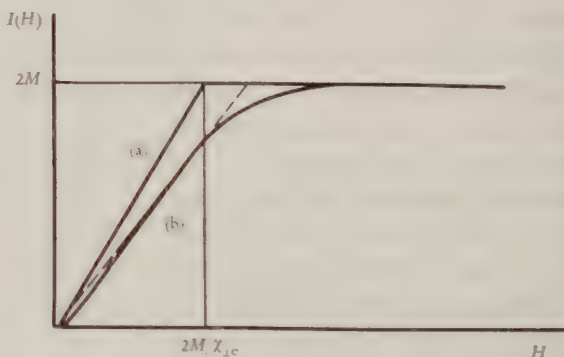


Fig. 6. curve (a): magnetization curve for field applied within the plane, where the anisotropy energy in the plane is completely neglected. curve (b): the powder magnetization curve, taking account of the anisotropy energy in and out of the plane.

* It is found, however, that χ_\perp is higher than the value given by (25) by one order. This discrepancy might be due to our inappropriate assumption of $\Delta E_s \gg \lambda_\perp$. The detailed analysis without this assumption will be given in another paper.

field which turns the magnetic moments in the plane and brings them to ferromagnetic saturation will depend on the direction of the field, one will expect the magnetization of the powder specimen as characterized by an initial increase corresponding to a certain average susceptibility, followed by a certain range corresponding to χ_{\perp} , and then by a gradual approach to ferromagnetic saturation (Fig. 6 (b)).* The experimental results for the powder specimens of CoCl_2 and NiCl_2 below their Néel points are qualitatively accounted for by this picture. We shall discuss these two crystals a little more in detail in the following lines.

In CoCl_2 , the orbital state of the free cobaltous ion, 4F , is split into three sublevels by the cubic field, the lowest sublevel being triply degenerate, and this lowest sublevel is further split into a doublet and a singlet by the trigonal field. We assume that the singlet is lower than the doublet; this assumption is not unreasonable, since it can be shown (see App. III) that the same trigonal field as that which makes the doublet lie lower than the singlet in FeCl_2 gives the reversed level scheme for CoCl_2 . However, the energy separation between the singlet and the doublet may not be large compared with the spin-orbit coupling energy. The combined action of the spin-orbit coupling and the trigonal field will split the spin-levels belonging to the singlet into two Kramers' doublets. Evidently, the separation between these two Kramers' doublets decreases with increasing orbital separation; with an orbital separation of 1000 cm^{-1} , the separation of Kramers' doublets will be about 100 K, and in the limit of vanishing orbital separation, it will be about 500°K. Thus we can safely assume that in the vicinity of the Néel temperature (25°K) only the lowest Kramers' doublet is populated.

Confining ourselves to the lowest Kramers' doublet, we can rewrite the Hamiltonian in terms of the effective spin operator s , of magnitude $1/2$, to which the true spin, S , is related by

$$S_z = \alpha_{\parallel} s_z, \quad S_x = \alpha_{\perp} s_x, \quad S_y = \alpha_{\perp} s_y.$$

In our orbital level scheme, α_{\parallel} is smaller than α_{\perp} and in the limit of large orbital separation, $\alpha_{\parallel} = 1$ and $\alpha_{\perp} = 2$.** Thus the exchange interaction, $2\mathbf{S}_i \cdot \mathbf{S}_j$, (either intra-layer interaction with $J < 0$ or inter-layer interaction with $J > 0$) is rewritten by s_i and s_j as

$$2J\{\alpha_{\parallel}^2 s_{xi} s_{xj} + \alpha_{\perp}^2 (s_{xi} s_{xj} + s_{yi} s_{yj})\}. \quad (41)$$

From (41) we see that the exchange energy is highly anisotropic in the lowest doublet, so that the spins orient perpendicularly to the c -axis in the antiferromagnetic state in

* The initial average susceptibility will be given by $1/3(\chi_{\parallel} + \chi_{\perp} + \chi_{\parallel c})$, where χ_{\parallel} and χ_{\perp} are the ordinary parallel and perpendicular susceptibilities, corresponding to a field applied in the plane, and $\chi_{\parallel c}$ is the susceptibility for a field applied parallel to the c -axis; we expect $\chi_{\parallel c}$ to be substantially smaller than χ_{\perp} if there is a large out-of-plane anisotropy energy, as in the case of CoCl_2 , but otherwise $\chi_{\parallel c} = \chi_{\perp}$. The second stage of magnetization will correspond to $(2/3)\chi_{\perp} + (1/3)\chi_{\parallel c}$ if $\chi_{\parallel c}$ is not equal to χ_{\perp} .

** In the limit of large orbital separation, the lowest doublet is specified by $S_z = \pm 1/2$, as one will easily understand by considering that the effective Hamiltonian of the second-order energy of the spin-orbit coupling is given by DS_z^2 , with positive D .

which the spins in the same layer point parallel and those in adjacent layers antiparallel.

The single crystal measurements by Bizette et al.² show that there is a large anisotropy of the susceptibility, i. e., below and above the Néel temperature. In fact, χ_{\parallel} is about seven times larger than χ_{\perp} , below the Néel temperature. We can understand this large anisotropy by considering that in antiferromagnetic state χ_{\parallel} is related to a weak inter-layer antiferromagnetic interaction and χ_{\perp} mostly to the anisotropy of the strong intra-layer ferromagnetic interaction given by (41). Also, a possible anisotropy of the g -factor might play a role.

The transition field to ferromagnetic state is given by $H = 2M/\chi_{\perp}$ when the field is applied within the plane. If we take $\chi_{\perp} = 0.4$ after Bizette,² supposing his χ_{\parallel} to be χ_{\perp} , and assume $3/\mu_B$ per ion for M , we obtain $H_c = 45,000$ o at absolute zero. With increasing temperature, H_c will decrease as the sublattice magnetization will decrease. This prediction is in accordance with the observation on powder specimen³ which shows that at 20.4°K the magnetization begins to saturate at about 28,000 o (see Fig. 2). For the field applied along the trigonal axis, H_c is expected to be much higher.

In NiCl_2 the anisotropy energy is expected to be smaller than in CoCl_2 . The single crystal data show no appreciable anisotropy of the susceptibility. This suggests that the anisotropy energy is much smaller than the inter-layer antiferromagnetic interaction whose order of magnitude is estimated from the relation $A = 1/\chi_{\parallel}$ and is $2J_1\tau_2 = 7 \text{ cm}^{-1}$. The main origin of the anisotropy energy must be the magnetic dipole interaction and the second order perturbation energy of the spin-orbit coupling, since the ground orbital state of Ni^{2+} is non-degenerate already in cubic field. The dipolar interaction will favour the plane perpendicular to the trigonal axis, with an energy difference of about 0.5 cm^{-1} between the axis and the plane. To determine the second part of the anisotropy energy, a detailed knowledge of the crystalline field would be required.

If we apply a field, the sublattice magnetizations of NiCl_2 will turn in the basal plane and make an angle with each other. If the field, which is applied with an angle to the plane, is increased, the sublattice magnetization will come out of the plane and at a field of $H_c = 2M/\chi_{\perp}$ transition to ferromagnetic state will take place. Using Bizette's value of $\chi_{\perp} \sim 0.1$, we obtain $H_c = 100,000$ o at absolute zero. The powder susceptibility data³ are not inconsistent with these predictions. However, Bizette's single crystal susceptibility shows a peculiar behaviour: it is field-dependent in the vicinity of the Néel temperature of 50 K, shows a minimum at 33 K, and below 33 K it ceases to be field-dependent. Both the powder and single crystal measurements give a positive paramagnetic Curie temperature, θ , of 68°K or 67°K for the susceptibility near 300 K, which could be understood qualitatively if one remembers that the major exchange interactions function ferromagnetically between ions in the same layer. Since, however, $\theta < T_N$, the simple molecular field approximation which always predicts $\theta > T_N$, will not apply.

Finally, it might be noted that the ferromagnetic interaction acting among the cations in the same hexagonal layer in all these chlorides can be understood by considering possible superexchange mechanisms. The author will deal with this topic in another paper.

The author would like to express his sincere thanks to Professor T. Nagamiya for

many valuable discussions. Thanks are also due to Professor K. Yosida and Professor S. Yomosa for their kind discussions and to Mr. Y. Yamada and Mr. H. Suzuki for the calculation of the crystalline field given in Appendix III. The present work was supported in part by a Grant in Aid from the Education Ministry.

Appendix I. Matrix elements of L

The matrix elements of L_z with respect to ψ defined in § 2 are obtained as

$$\begin{aligned}(g1|L_z|g1) &= -(g2|L_z|g2) = 1 + (3/2)(\cos 2\vartheta - (1/3)) \\ (b1|L_z|g1) &= (b2|L_z|g2) = -(3/2)\sin 2\vartheta \\ (b1|L_z|b1) &= -(b2|L_z|b2) = 1 - (g1|L_z|g1),\end{aligned}\tag{A.1}$$

other elements being equal to zero. The matrices of L_x and L_y have no diagonal elements. Defining $L^\pm = L_x \pm iL_y$, the matrix elements of L^\pm between the lowest doublet and other states are given by

$$\begin{aligned}(g2|L^\pm|g1) &= 0, \quad (s|L^+|g1) = (s|L^-|g2) = -\sqrt{6}\sin\vartheta, \\ (b1|L^+|g2) &= (b2|L^-|g1) = 2.\end{aligned}\tag{A.2}$$

Appendix II. Evaluation of Φ_1 and Φ_2

First we discuss the approximate value of Φ_{2z} . As defined in § 3, Φ_{2z} is the coefficient of the inter-sublattice dipole interaction. In FeCl_2 , the layers of cations are well separated from each other by the layers of anions, so that we can assume safely that the magnetic dipoles in the different sublattice are continuously distributed in each layer. The continuous dipole layer which extends infinitely does not produce a magnetic field outside the layer. This means

$$\Phi_{2z}M + (4\pi/3)M - 4\pi M = 0.$$

Thus we obtain

$$\Phi_{2z} = 8\pi/3 = 8.37.\tag{A.3}$$

Actually Suzuki⁽²¹⁾ obtained $\Phi_{2z} = 8.46$ for MnBr_2 by a rigorous calculation. In MnBr_2 , the separation between the adjacent cation layers is 6.19 Å and the distance between the nearest-neighbouring Mn ions within a layer is 3.82 Å, while in FeCl_2 the corresponding distances are 5.85 Å and 3.58 Å, the ratios coinciding within 1 per cent. Thus the error in (A.3) will be within 2 per cent. For Φ_{1z} we separate the contribution from the dipoles within the same layer from those from other layers. For the latter we obtain, by the same consideration, a value $8\pi/3$. The former has been calculated by Suzuki⁽²¹⁾ as -30.2 for MnBr_2 . Considering the difference of the lattice constants, we obtain the corresponding value for FeCl_2 as -30.3 . Thus Φ_{1z} is obtained as

$$\Phi_{1z} = -21.9.\tag{A.4}$$

The error in (A.4) will be smaller than that in (A.3). The x, y components of ψ can be obtained from (A.3) and (A.4) through the relation, $\psi_x = \psi_y = -(1/2)\psi_z$.

Appendix III. Calculation of the crystalline field

According to the paramagnetic resonance experiment on NiCl_2 , the g -value of the nickelous ion is found⁽²²⁾ to be $g=2.25$. Since nickelous ion has a non-degenerate orbital level in the cubic field, its g -factor can be expressed approximately as

$$g=2+8|\lambda|/\Delta E, \quad (\text{A.5})$$

where ΔE is the separation between the ground orbital level and the first excited orbital level in the cubic field. From $g=2.25$ and (A.5) we see that the separation of the orbital levels due to the cubic field is of the order of 1 eV., or 10^4 cm^{-1} ($\lambda=240 \text{ cm}^{-1}$).

It is difficult to calculate the cubic field, the major part of which comes from the chlorine ions in contact (see ref. 4) for the discussion of the crystalline field). Therefore an attempt was made to calculate only the trigonal part of the crystalline field. We can infer from this part an approximate orbital level scheme, assuming that the splitting due to the cubic field is about 1 eV.

In FeCl_2 , we may neglect the trigonal field arising from a small departure from the cubic arrangement of chlorine ions. The main trigonal field will be contributed from the surrounding cations and the induced polarizations of the chlorine ions. For the latter we assume a point dipole moment of magnitude p pointing in the direction of the c -axis. The potential arising from these sources can be expressed as

$$V = (C_{20} + pC_{20}')r^2P_2(\cos\vartheta) + (C_{40} + pC_{40}')r^4P_4(\cos\vartheta) \\ + (C_{43} + pC_{43}')r^4P_4(\cos\vartheta)\cos 3\varphi. \quad (\text{A.6})$$

We calculated C_{20} , C_{20}' , etc., by the direct sum method.⁽²⁰⁾ The results are listed in Table I.

To determine p , we calculate the electric field, E , at each chlorine site. This is expressed as

$$E = E_0 + pE', \quad (\text{A.7})$$

so that

$$p = \alpha(E_0 + pE') \quad \text{or} \quad p = \alpha E_0 / (1 - \alpha E'), \quad (\text{A.8})$$

where α is the polarizability of chlorine ion. Assuming $\alpha = 2.97 \text{ \AA}^3$, we obtain $p = 0.751 e\text{\AA}$, where e is the elementary charge.

Since the Hartree or Hartree-Fock function is not available for Fe^{2+} , we calculated the average of r^2 and r^4 using the Hartree-Fock function for Mn^{2+} ⁽²³⁾ and also the Hartree function for Co^{2+} .⁽²⁵⁾ The former gives $\langle r^2 \rangle = 0.41 \text{ \AA}^2$ and $\langle r^4 \rangle = 0.375 \text{ \AA}^4$ and the latter $\langle r^2 \rangle = 0.455 \text{ \AA}^2$ and $\langle r^4 \rangle = 0.53 \text{ \AA}^4$. If we use alternatively the Slater function for Fe^{2+} , we obtain $\langle r^2 \rangle = 0.90 \text{ \AA}^2$ and $\langle r^4 \rangle = 1.31 \text{ \AA}^4$. The maximum point of the radial density

of d -electrons of Mn^{2+} is more distant from the nucleus in the Slater function than in the Hartree-Fock function. However, according to the result of neutron diffraction experiment for Mn^{2+} , the maximum point lies between the value given by the Slater function and that given by the Hartree-Fock function. Considering these facts, we assumed arbitrarily $\langle r^2 \rangle = \langle r^4 \rangle = 0.45$ (in Å unit) in our calculation.

We find that the potential determined in this way makes the orbital singlet lower than the doublet. The results for ΔE_s , ΔE_h , ϑ and $\langle L \rangle$ for three values of p are listed in Table II. It is found that for p -values smaller than about 0.5 eÅ the doublet comes out to be lower than the singlet. This result does not depend sensitively on the choice of the values of $\langle r^2 \rangle$ and $\langle r^4 \rangle$, because the sign of $C_{20} + pC_{20}'$ is mainly responsible to determine the sign of ΔE_s .

Smaller p -values are not unreasonable, since one would expect that the overlap of the charge clouds of the cation and anion will diminish the potential at the cation arising from the chlorine dipole moments. Also the repulsion between the electron clouds will tend to diminish the polarization of the chlorine ion.

In the case of CoCl_2 , the orbital levels were calculated by using the same potential with $p=0$. In this way the desired orbital level scheme is obtained and the separation between the singlet and the doublet turns out to be about 1000 cm^{-1} . For certain large p -values, the level scheme is reversed.

Table I

C_{20}	$0.1105 \text{ e}^2/\text{Å}^3$	C_{20}'	$-0.210,0 \text{ e}/\text{Å}^4$
C_{40}	$-0.007,26 \text{ e}^2/\text{Å}^5$	C_{40}'	$-0.009,45 \text{ e}/\text{Å}^6$
C_{43}	$0.000,049 \text{ e}^2/\text{Å}^5$	C_{43}'	$-0.001,033 \text{ e}/\text{Å}^6$
E	$0.418,9 \text{ e}/\text{Å}^2$	E'	$-0.221,2 \text{ 1}/\text{Å}^3$

used lattice constants: $a=6.20 \text{ Å}$, $\alpha=33^\circ 33'$, $u=0.25$.

Table II

	$p=0$	$p=0.37$	$p=0.74$
$\Delta E_h (\text{cm}^{-1})$	9223	8017	7032
ΔE_s	2540	420	-1589
L	1.351	1.162	0.892
$\cos \vartheta$	0.8853	0.849	0.794
$\sin \vartheta$	0.465	0.529	0.608

References

- 1) H. Bizette, *Ann. de Phys.* (12) **1** (1946), 316; Thesis (1946), 84; *Jour. de Phys. et le Rad.* **12** (1951), 161.
- 2) C. Starr, F. Bitter and A. R. Kaufmann, *Phys. Rev.* **58** (1940), 977.
- 3) T. Nagamiya, K. Yosida and R. Kubo, *Advances in Phys.* **4** (1955), 1.
- 4) J. Kanamori, *Prog. Theor. Phys.* **17** (1957), 177.
- 5) L. Landau, *Phys. Z. Sowjet.* **4** (1933), 675.
- 6) R. W. G. Wyckoff, *Crystal Structures*, Interscience Pub. N. Y. (1948).
- 7) M. K. Wilkinson and J. W. Cable, *Bull. Amer. Phys. Soc.* **1** (1956), 190.
- 8) O. Trapeznikowa and L. Schubnikow, *Physik Z. Sowjet.* **7** (1935), 66.
- 9) H. Bizette, C. Terrier and B. Tsai, *Comptes Rendus* **243** (1956), 895, 1295.
- 10) L. Néel, *Comptes Rendus*, **242** (1956), 1549.
- 11) For instance, Bleaney and Stevens, *Réports on Prog. in Phys.* **16** (1953), 108.
- 12) M. H. L. Pryce, *Proc. Phys. Soc. (London)* **63** (1950), 25.
- 13) J. Kanamori, K. Motizuki and K. Yosida, *Busseiron Kenkyu* **63** (1953), 28; K. Yosida, *Ann. Rep. Sci. Fac. Osaka Univ.* **1** (1953), 19.
- 14) C. J. Gorter and T. van Peski-Tinbergen, *Physica* **22** (1956), 273.
- 15) K. Honda and T. Ishiware, *Science Rep. Tohoku Univ.* **4** (1915), 215.
- 16) H. R. Woltjer and E. C. Wiersma, *Proc. Roy. Acad. Amsterdam* **32** (1929), 736.
- 17) see ref. 1); M. Foex, *Ann. de Phys.* (9) **16** (1921), 255.
- 18) W. C. Koehler and E. Q. Wollan, *Bull. Amer. Phys. Soc.* **1** (1956), 190.
- 19) H. Bizette, *Comptes Rendus* **245** (1957), 507.
- 20) C. G. B. Garrett, *J. C. P.* **19** (1951), 1154.
- 21) H. Suzuki, private communication.
- 22) J. W. Leech and H. T. Manuel, *Proc. Phys. Soc. (London)* **B69** (1956), 210.
- 23) J. R. Tessman, A. H. Kahn and W. Shockley, *Phys. Rev.* **92** (1953), 890.
- 24) D. R. Hartree, *Proc. Camb. Phil. Soc.* **51** (1954), 126.
- 25) M. Ariyama, K. Kayama and M. Sugimoto, *J. Phys. Soc. Japan* **12** (1957), 285.
- 26) J. Kanamori, T. Moriya, K. Motizuki and T. Nagamiya, *J. Phys. Soc. Japan* **10** (1955), 93.

Mass Reversal and Weak Interactions*

Tadashi OUCHI**

Department of Physics, University of Maryland, College Park, Maryland

(Received July 2, 1958)

It is shown that the special mass reversal MR(1) and MR(2) can be applied to both the weak Fermi type interaction and the weak Yukawa type interaction without any internal contradiction in the theory of MR(4) and MR(3) (for strong interactions). The selection rules obtained are consistent with experiment and suggest taking the universal Fermi interactions with $V-A$ coupling as the primary interactions for all the lepton processes. The mass reversibility of the theory under MR(1) requires that the space parity is not conserved in all the weak interactions and that the mass of the neutrino must be exactly zero. MR(2) invariance allows only a linear combination of (AV) couplings when the coupling is independent of the explicit masses. It is assumed that leptons are conserved and are C or D type fields under mass reversal, guaranteeing the conservation of baryons automatically. From the requirement of MR(4) and the consistency of the theory, the neutrino is shown to be a mass parity doublet. It is shown from the application of mass reversal to the lowest order diagrams in g that the unwanted lepton processes are forbidden or explained even when the μ^- , e^- and ν are taken as the particles. The prediction of the rarity of $\pi \rightarrow e + \nu$, $\pi \rightarrow e + \nu + \gamma$ and $K \rightarrow e + \nu$ is given by extending the universal interaction proposed by Huang and Low in such a way that it is invariant under mass reversal. If this is true, the neutrino must have four components in a strict sense, and in the usual experiments it is observed as if it were a two component field. If MR(3) is assumed to be applied even to Fermi interactions, MR(3) requires uniquely that the universal interactions have $V-A$ coupling, and the helicity of the neutrino is -1 . The decays $K \rightarrow (\pi) + \bar{\nu} + \nu'$ are shown to be ruled out by the mass parity conjugation invariance and MR(1), provided the leptons are C or D type fields under mass reversal. The mass parity is shown to be related closely to the strangeness which is the important difference between our theory and Sakurai's one.

§ 1. Introduction

It is pointed out in a previous paper¹⁾ that mass reversal is a kind of spinor transformation and is independent of charge conjugation, space reflection or time reversal. Thus special mass reversals (for example, MR(1) and MR(2)) can be considered consistently with usual mass reversal MR(4), which means that particles for which mass reversal is applied are in principle chosen arbitrarily. Hereafter we shall denote the operation of mass reversal simply by MR(n) where n is the number of particles whose masses are being reversed. The weak interactions of Fermi type are taken in this paper as the most

*) A preliminary paper on *Mass Reversal and Lepton processes* was circulated in December, 1957, which was the revised version of the report in collaboration with K. Senba and M. Yonezawa which appeared in *Soryushiron-Kenkyu* (mimeographed circular in Japanese) 14 (1957), 618. And. See, also, T. Ouchi, K. Senba and Y. Yonezawa, *Nuovo Cimento* 8 (1958), 708.

**) On leave of absence from Hiroshima University.

probable primary interactions responsible for the lepton processes. It is one of the main problems to examine whether these special mass reversals MR(1), MR(2) and the usual MR(4) can really provide the useful selection rules for the lepton processes in a consistent way with both strong and weak interactions in the baryon-meson system.

An attempt very similar to these special mass reversals was recently proposed by Nishijima²⁾ to derive the selection rules for lepton processes. However, we claim that the assumption of invariance under mass reversals MR(1) and MR(2) is less restrictive than Nishijima's assumptions to produce essentially the same results. The meaning of mass reversal will be more clear and the assumptions in this paper will be justified once the structure of the elementary particles is established.

The main results obtained by applying MR(1), MR(2) and MR(4) for leptons and baryons are the following :

From the requirement of mass reversal MR(1) invariance, it is concluded that the space parity is not conserved in all the weak interactions and the mass of neutrinos must be exactly zero. Consequently the Fierz interference term in β -decay and the magnetic moment of the neutrino must be exactly zero. From the MR(2) it is shown that a linear combination of (AV) couplings is allowed when the couplings do not include the explicit masses and that a linear combination of (STP) couplings is possible when the couplings include the explicit masses in the proper way (for example eq. (18) is the case.) From MR(4) and the consistency of the theory, the neutrino is shown to be a mass parity doublet. Then the leptons are assumed to be conserved and to be C or D type fields under mass reversal to guarantee the conservation of baryons. By taking the μ^- , e^- , ν and ν' as the particles and the Fermi interactions with $V-A$ couplings, Sachs' assumption that all the lepton processes must involve at least one neutrino is essentially proved. The decay modes such as $K \rightarrow \pi + \mu^- + \mu^+$, $K \rightarrow \pi + e^- + e^+$ which are not forbidden by Konopinski-Mahmoud's³⁾ procedure are forbidden from the requirement of MR(4). For the prediction of the rarity of the modes $\pi \rightarrow e + \nu + (\bar{\nu})$ the universal $V-A$ interaction was modified so as to be invariant under mass reversal. The new terms needed by this modification are shown to have no contribution for those processes, so that these new interactions are reduced to Huang-Low's⁴⁾ universal interaction. Consequently the $\pi \rightarrow e + \nu$, etc., are predicted in the same way as Huang and Low. If this is taken seriously, the neutrino is described by a four component field in a strict sense and only the proton and electron are always left-hand polarized. This is probably due to the fact that both electron and proton are stable. MR(3) was defined to forbid the mode $\mu \rightarrow e + e^- + e^+$ and seems to imply that the universal Fermi interactions have the $V-A$ couplings and the helicity of the neutrino is -1 which is consistent with the latest experiment⁵⁾. The decays $K \rightarrow (\pi) + \bar{\nu} + \nu'$ are shown to be ruled out by the mass parity conjugation invariance and MR(1) in $V-A$ theory provided the leptons are C or D type field under mass reversal.

The most important points to be emphasized are that our mass reversal is applied both to strong and weak interactions and the mass parities are related closely to the strangeness and that the weak interactions for baryon and meson systems are the universal Yukawa type interactions with gradient coupling. These are the essential differences

between our mass reversal and Sakurai-Hori-Wakasa's⁽⁴⁾ theory. The comparison of the prediction of the universal $V-A$ interactions was not presented in this paper in order to avoid the overlap with those papers by Feynman-Gell-Mann⁽⁷⁾ and by Sudarshan-Marshak⁽⁸⁾ and by Sakurai.

§ 2. Outline of the theory of mass reversal

In order to clarify the meaning of mass reversal, it will be useful to recall Nishijima's theory⁽²⁾ for leptons. He found that the following two sets of equations (I) and (II) yield the same S -matrix, but not the same Hamiltonian in quantum electrodynamics.

$$\begin{aligned} \{\gamma_\mu(\partial_\mu - ieA_\mu) + m\}\psi &= 0 \\ \square A_\mu &= -ie\bar{\psi}\gamma_\mu\psi \end{aligned} \quad (\text{I})$$

$$\begin{aligned} \{\gamma_\mu(\partial_\mu - ieA_\mu) + me^{2i\alpha\gamma_5}\}\psi &= 0 \\ \square A_\mu &= -ie\bar{\psi}\gamma_\mu\psi \\ (\alpha : \text{real constant}), \end{aligned} \quad (\text{II})$$

This fact is due to the special property of γ_μ that γ_μ can absorb the transformation factor $e^{-i\alpha\gamma_5}$ appearing in the physical quantities in such a way that $e^{-i\alpha\gamma_5}\gamma_\mu e^{-i\alpha\gamma_5} = \gamma_\mu$. Next he assumed that the two sets of equations above mentioned yield the same S -matrix even in the weak lepton processes, and derived several interesting results which seem to be consistent with experiment.

Although the essential point of the Nishijima theory is that α is arbitrary, we may fix α to $\pi/2$. Then his attempt is equivalent to the concept of our mass reversal. This will be the most simple explanation of the mass reversal is less restrictive than his theory in the sense that α is fixed. In fact the theory of mass reversal can be developed more generally in a similar way as space reflection.

In q -number theory of mass reversal, we have shown in the previous paper⁽¹⁾ that mass reversal is equivalent to a kind of spinor transformation and, therefore, can be applied even to the neutrino with zero mass. The arguments can be formulated in the interaction representation as follows:

The mass-reversibility is to assure for S -matrix the relation

$$\begin{aligned} \langle f|S|i\rangle &\equiv \langle f; \kappa|S[\kappa]|i; \kappa\rangle \\ &= \langle f; -\kappa|S[-\kappa]|i; -\kappa\rangle \\ &\equiv \langle f'|S'|i'\rangle \end{aligned} \quad (1)$$

where $i(i')$ and $f(f')$ represent the initial and final states respectively and S' indicates that it is constructed from a set of equations of type (II) in the same way as S is made from set (I). If there exists a time-independent unitary operator M for mass reversal satisfying $|i'\rangle \equiv |i; -\kappa\rangle = M|i\rangle$, eq. (1) is translated into

$$S' = MS[\kappa]M^{-1} \quad (2)$$

or in terms of interaction Hamiltonian H ,

$$H' \equiv H[-\kappa] = \rho_H (MH(\kappa)M^{-1}). \quad (3)$$

And in general for any physical quantities Q 's we have

$$Q' \equiv Q[-\kappa] = \rho (MQ(\kappa)M^{-1}) \quad (4)$$

where ρ is a sign function for Q under mass reversal. In usual case ρ_H must be taken as $+1$ as H is energy. $H(\kappa)$ contains, in general, a certain number of explicit masses κ_i and fields $\psi_i(\kappa)$. Then $H(-\kappa)$ has the same form as $H(\kappa)$, but with explicit masses $-\kappa_i$ and fields $\psi_i(-\kappa)$ instead of $\psi_i(\kappa)$. Let us take as the most general interactions of Fermi type

$$H = \sum_{\alpha} f_{\alpha} (\bar{\psi}^a O_{\alpha} \psi^b) (\bar{\psi}^{b'c} O_{\alpha} \psi^{a'}) \quad (5)$$

where the order of fermions is fixed for simplicity so that both a and b are baryons or leptons, and take into account the special mass reversal MR(1) and MR(2) together with the usual MR(4). This possibility was pointed out in the previous paper^{*} from the physical meaning of mass reversal. As the double MR(2) is in a certain case equal to MR(4) and double MR(1) is in a certain case equivalent to MR(2)^{*}, it will be a sufficient extension to assume the value of ρ_H for MR(2) and MR(1) to be (± 1) and $(\pm 1, \pm i)$ respectively. Corresponding to this, eq. (1) should be modified slightly. The meaning of these assumptions will be clarified in the next section.

As was shown in the previous paper^{*}, the operator M exists and the spinor fields ψ_i the boson ϕ_i transform as follows:

$$\begin{aligned} M\psi_i(x, \kappa_i)M^{-1} &= \omega_i \gamma_5 \psi_i(x, -\kappa_i), \\ M\phi_i(x, \kappa_i)M^{-1} &= \omega_i' \phi_i(x, -\kappa_i), \end{aligned} \quad (6)$$

where ω_i is the relative phase of ψ_i under mass reversal and takes the value $+1$, $+i$ and ω_i' is the mass parity of bosons, taking the value $+1$. Now we can derive the selection rules from the requirements of mass reversibility of the Hamiltonian. Assuming that the f_{α} do not involve any explicit mass, we get from eqs. (3), (5) and (6),

$$\omega_a^* \omega_b \omega_c^* \omega_d = 1, \quad \text{from MR(4)} \quad (7a)$$

$$\pm \omega_c^* \omega_d = \alpha \quad \text{with } \alpha = \pm 1, \text{ from MR(2)} \quad (7b)$$

where we take the $+$ sign for O_{α} with V or A coupling and the $-$ sign for O_{α} with S , T or P coupling.

$$O_{\alpha} \gamma_5 \omega_d = \beta O_{\alpha} \quad \text{with } \beta = \pm 1, \pm i \text{ from MR(1)} \quad (7c)$$

As ω_d is limited to the value ± 1 and $\pm i$, eq. (7c) means

^{*}) This situation will be more clearly understood from the following fact. The operator M for MR(4) is given by $M = M_a M_b M_c M_d$ where M_i is the solution of eq. (6) for M and the mass reversal operator for MR(1) applied to the i -particle. Therefore the double MR(1), i. e. the product of MR(1) with M_a and MR(1) with M_b is equal to MR(2) with $M_a M_b$, and so on.

$$O_a = O_a'(1 \pm \gamma_5) \text{ or } O = O_a'(1 \mp i\gamma_5) \quad (8)$$

where the case $O_a = O_a'(1 \mp i\gamma_5)$ is forbidden if the time reversal invariance is preserved.

(7b) means that there occurs either a linear combination of S , T and P couplings or a linear combination of V and A couplings, which are the same results obtained by Tiomno⁹⁾. (8) means that space parity is not conserved in the weak Fermi interactions.

It will be noted that MR(2) is a special case of double MR(1) and should be applied to lepton pair or baryon pair, and that the relationship β_{at} (for STP couplings) $= -\beta_{at}$ (for AV couplings) is assumed in the following discussions because eq. (7c) shows that β_{at} will depend on the type of O_a .

§ 3. Application of various mass reversals to leptons

Now we are ready to apply all kinds of mass reversal to Fermi type interactions responsible for lepton processes. As the form invariance of the theory under MR(1), MR(2) and MR(4) simultaneously is required, eq. (1) should be replaced correspondingly by

$$|\langle f; \kappa | S[\kappa] | i; \kappa \rangle|^2 = |\langle f; -\kappa | S[-\kappa] | i; -\kappa \rangle|^2 \quad (9)$$

where $|\dots|^2$ implies that one takes care of the average over the initial spin states and the summation over the final spin states. The above change is necessary because α and β for MR(2) and MR(1) are not always equal to $+1$.

To assure the consistency of the theory of mass reversal, the mass parity assignment for hyperons and mesons is necessary. We have formerly¹⁰⁾ proposed the form of the strong interactions and the universal weak interactions for the hyperon and meson system. As the most probable primary interactions of the Yukawa type for this system, we suggested

$$G \bar{\psi}_a \gamma_5' \psi_b \phi_c \text{ for the strong interactions} \quad (10)$$

and

$$g \bar{\psi}_a \gamma_5'' \gamma_\mu \psi_b \phi_c \text{ for the weak interactions,} \quad (11)$$

where G and g are assumed to be natural constants¹¹⁾ with the magnitude of order $1 \sim 10$ and $10^{-13} \sim 10^{-14}$ in units of $\hbar = c = \mu$ (π meson mass) similar to e for the electromagnetic interaction and γ_5' is taken to be either γ_5 or 1 if the space parity is conserved in the weak interactions, but all γ_5'' can be equal to $1 \pm \gamma_5$ if the space parity is not conserved in the weak interactions. Assigning the mass parity as follows:

$$\Xi, N \in A, \sum, A \in B, \pi \in -1, K \in +1, \\ A_\mu \in +1. \quad (12)$$

where \in means that Ξ, N belong to the fields of type $A^{(10)}$ under mass reversal, etc., we can derive good selection rules which are consistent with experiment. If we connect the mass parity with the attribute $d^{(2)}$ so that $\omega = (-1)^a$ for fermions and $\omega = -(-1)^a$ for bosons, our selection rule for hyperon and meson system reads from MR(3) with $\rho_H = +1$ in (3)

$$(-1)^{\sum_i a_i} = (-1)^{\sum_f a_f} (-1)^{n.W.} \quad (13)$$

where i and f indicate the initial and final states, the summation should be done over all the particles present in the initial and final states and $n.W.$ means the number of vertices with weak coupling g in the Feynman diagram. From (13), we can derive the even-odd rule,

$$\begin{aligned} \Delta a &= 0 \pmod{2}, \text{ for the strong interactions} \\ \Delta a &= 1 \pmod{2}, \text{ for the weak interactions.} \end{aligned} \quad (14)$$

But eq. (13) can be taken more naturally to imply that Sachs' rule can be obtained if a 's are chosen in such a way that decay $\Xi^- \rightarrow N + \pi^-$ is forbidden

$$\Delta a = n \equiv n.W. \quad (15)$$

Although Sachs assumed that the transition having the larger n becomes much weaker, the strength of the transitions having $\Delta a = n$ can be estimated in our theory to be of order g^n , therefore very weak unless $n=0$ or 1. This is the reason why there exist the strong and weak interactions observed in the hyperon-meson system.*

In order to discuss all the weak interactions in a unified manner, the knowledge about α and β is needed from the interactions established for the baryon and meson system. As is easily shown, the weak interactions (11) with $1 \pm \gamma_5$ for γ_5 's are invariant under MR(1) and MR(2) with $\alpha = -1$. Consequently we assume in general the following relations:

$\alpha = +1$, when $\sum_i s_i = \text{even}$, neutrino ν is produced for the lepton process

$\alpha = -1$, when $\sum_i s_i = \text{odd}$, neutrino ν' is produced from the lepton process, (16)

where the summation runs over the baryons contained in the interaction and s represents the strangeness of the particle. (That the leptons have zero strangeness is assumed.) This will also be natural from the point of view that the strangeness of the nucleon is zero and the strangeness is related to the mass parity and that the strong interactions satisfy under MR(2) the similar relationship as (16) which has the opposite sign for α . Now let us apply MR(4), MR(2) and MR(1) to the Fermi interactions $n \rightarrow p + e + \nu$ and $\Lambda \rightarrow p + e + \nu'$ as the typical examples. Then we have from eqs. (7a), (7b) and (7c), provided the coupling constants do not contain explicit masses,

$$\omega_p^* \omega_N \omega_e^* \omega_\nu = 1, \quad \omega_p^* \omega_\Lambda \omega_e^* \omega_{\nu'} = 1 \quad (17a)$$

and $O_\alpha = \gamma_\mu (1 \pm \gamma_5)$ for all the Fermi interactions (17b)

where neutrino ν in β -decay is distinguished from the neutrino in Λ -decay. (17b) implies that the universal Fermi interaction is possible between baryon and lepton families and

* Similarly it is derived that the lepton processes having $\Delta a = 0$ are allowed and their strengths of the transition are of order g , provided that they proceed through the interactions (10) and (5).

between lepton and lepton families. And $\omega_\nu = -\omega_{\nu'}$ is deduced from eqs. (12) and (17a).

In order to obtain definite solutions for ω 's from (17a) it is assumed that leptons are conserved and all the leptons belong to C or D type fields⁽¹⁰⁾ under mass reversal. Then $\omega_\mu = \omega_e = \omega_\nu \in C$ and $\omega_{\nu'} \in D$ (or $\omega_\mu = \omega_e = \omega_\nu \in D$ and $\omega_{\nu'} \in C$) is the solution and the conservation of baryons is preserved in the same way as treated by Yang and Tiomno⁽¹³⁾. If we apply MR(1) to baryons in order to obtain the knowledge about β we find the relationship between β and ω so that $\beta_a = \beta_b$ if $\omega_a = \omega_b$ due to $O_a = \gamma_\mu (1 \pm \gamma_5)$. This relation suggests indirectly that (16) was a good choice. By using this relationship, both ν and ν' are shown to have the same helicity which means that neutrino is a mass parity doublet⁽¹⁴⁾. This strange situation is not the defect of the present theory. Rather it originates only from the formal ambiguity of mass reversal for the neutrino due to its vanishing rest mass.

It is easily shown by using the same knowledge about α and β such as (16), and β_u (for AV couplings) $= -\beta_d$ (for STP couplings), etc., that the following β -decay interaction with the coupling constants involving the explicit masses is invariant under MR(1), MR(2) and MR(4):

$$G(\bar{\psi}_p \gamma_\mu (1 \pm \gamma_5) \psi_{N(\Lambda)}) (\bar{\psi}_e \gamma_\mu (1 \pm \gamma_5) \psi_{\nu(\nu')}) + G \frac{m_e}{m_p} \left\{ (\bar{\psi}_p (1 \mp \gamma_5) \psi_{N(\Lambda)}) (\bar{\psi}_e (1 \mp \gamma_5) \psi_{\nu(\nu')}) \right. \\ \left. + \text{tensor coupling} \right\}, \quad (18)$$

where m_e and m_p are the masses of electron and proton. This interaction implies that the proton and the electron are always left-hand (right-hand) polarized when the strong interactions are switched off and the small amount of right-handed (left-handed) neutrino can exist together with the normal left-handed (right-handed) neutrino or that the neutrino has four components in a strict sense but it is observed almost as the two-component particle⁽⁷⁾. Now we can prove from (9) that the mass of the neutrino must be exactly zero whenever this type of interaction is considered. Let us consider the interference term corresponding to the Fierz term which is proportional to

$$F_\mu G \frac{m_e}{m_p} S_p \{ (\gamma p + im_e) O'_\mu (\gamma' p' + im_\nu) \beta O''^\dagger \beta \} \quad (19)$$

where $O'_\mu = \gamma_\mu (1 \pm \gamma_5)$, $O'' = 1 \mp \gamma_5$ and F_μ is a vector made of nucleon wave functions. Then it turns out by a trivial calculation that this term is equal to $4i P_\mu F_\mu m_e m_\nu / m_p$, which is not invariant under MR(1) for the neutrinos unless $m_\nu = 0$. In the same way we can show that $m_{\nu'} = 0$. It should be pointed out that MR(1) invariance requires that the mass of the neutrino is exactly zero, provided that right-handed and left-handed neutrinos can coexist. Such an example was given by Nishijima⁽¹⁵⁾ to forbid the unwanted processes in the lepton processes. His results are equivalent to taking the μ^+ , e^- , ν as the particles and assuming the conservation of leptons which was the original idea of Konopinski-Mahmoud. The main difficulties of their theory are that the decay modes emitting the μ pair or electron pair such as $K \rightarrow \pi + \mu^+ + \mu^-$, $K \rightarrow \pi + e^+ + e^-$ cannot be forbidden, and that the experiments have shown that μ^- , e^- and ν are particles.

§ 4. Selection rule imposed by mass reversal

In this section we will show that mass reversal invariance predicts why the unwanted decay modes are not observed in the lepton processes when muon (μ^-), electron (e^-) and neutrinos (ν and ν') are taken as particles as proved by experiment. In this paper we take as the primary interactions the strong interactions (10), the universal weak interactions (11) and (5) with $V-A$ couplings and the electromagnetic interactions of the form as $e \bar{\psi}_a \gamma_\mu \psi_a A_\mu$. As a direct consequence of this, Pauli type electro-magnetic interactions having the form $(e/2m_\nu) \bar{\psi}_a \sigma_{\mu\nu} \psi_a F_{\mu\nu}$ only, are allowed from the requirement of mass reversal⁽¹⁰⁾. Thus the magnetic moment of the neutrino should be exactly zero because the neutrino has zero mass and the above form of interaction is impossible. The consideration of the lowest order diagrams with respect to the weak coupling constant g will be enough because g is very small.

First of all we show that the possible decay modes such as $K \rightarrow \pi + \mu^- + \mu^-$, $K \rightarrow \pi + e^+ + e^-$, $K \rightarrow \pi + \nu + \bar{\nu}$ are forbidden. The decay of the K -particle is understood as a result of the virtual process in which the K becomes a baryon loop which decays into the muon pair or the electron pair, etc., by the following

$$A \rightarrow N + \mu^+ + \mu^-, A \rightarrow N + e^+ + e^- \text{ and } A \rightarrow N + \nu + \bar{\nu}. \quad (20)$$

But these interactions are not invariant under MR(4) because the μ^- , e^- and ν are C type fields and A and N are the B and A type fields respectively. For the same reason the decay mode $K \rightarrow \mu + \mu_0$ (μ_0 is the neutral counter-part of muon) suggested by Marshak and Sudarshan¹⁷ is forbidden. The mode $K \rightarrow \pi^- + e^-$ does not conserve the leptons and therefore is forbidden. The decay mode $\mu^- \rightarrow e^- + \gamma$ can take place only in a diagram of the second order in g . Therefore this is hardly observed. The most important problem will be to explain the decay ratios $\mu^- \rightarrow (\pi^- + e^- + \nu) / (\pi^- + \nu + \bar{\nu}) = 10^{-5}$ and $\mu^- \rightarrow (\pi^- + e^- + \nu + \bar{\nu}) / (\pi^- + \mu + \nu) \approx 10^{-5}$. This is solved in the same calculation as Huang and Low if we adopt the interaction (18) which is a slight modification of the $V-A$ interactions. The difference between the interaction (18) and Huang-Low's interaction is that scalar and tensor couplings are not included in the latter. And fortunately the scalar coupling has no contribution for those decays and the small amount of tensor coupling does not necessitate any change for our purpose as was shown by them. In the same way it is understood, provided K -baryon strong interaction have γ_5 coupling, why the decay $K \rightarrow e^+ + \nu$ is not observed while $K \rightarrow \pi^0 + e^- + \nu$ is observed. Thus it seems to be a good assumption that the universal Fermi interactions have the form of (18) whenever both electron and proton take part in them. This was assumed for the present from the conjecture that the special interaction needed for the prediction of $\pi^- \rightarrow e^- + \nu$ might be closely related to the stability of the proton and electron. Then we can forbid the capture mode $\mu^- + p \rightarrow p + e^-$. For this case the interaction of the form (18) will work and then the mass of the μ is required to be zero from the MR(1). Unfortunately there exists no charged particle with zero mass. Thus it is impossible. We argue that the decay mode $\mu^- \rightarrow e^- + e^- + e^-$ will be forbidden for the following reasons: The mass reversal MR(1) for the electron is in this case just triple MR(1) which will be defined to be MR(3) for the Fermi interaction from

the same reasoning as that MR(2) was the special case of the double MR(1). The other MR(3) is known to be applied to the Yukawa type interactions and requires that $\rho_H = +1$. Then it will be natural to take $\rho_H = +1$ for all MR(3). Due to this assumption the above process is forbidden because the e^- is a C-type field, therefore $\rho_H = +i$. To check this assumption, let us apply the same argument to the interaction $(A O_\alpha N) (\bar{N} O_\alpha N)$ in Sakata's composite particle model¹⁶⁾ for baryons and mesons which corresponds to the decay $\Lambda^0 \rightarrow N + \pi^0$. The result is that O_α should be equal to $\gamma_\mu (1 + \gamma_5)$, which means that the universal Fermi interaction has $V-A$ coupling, and the helicity of neutrino is -1 . These are consistent with experiment³⁾, which shows that the definition of MR(3) is correct.

Finally we will show that the decays $K^0 \rightarrow \bar{\nu} + \nu'$ and $K \rightarrow \pi + \bar{\nu} + \nu'$ are forbidden from the invariance by mass parity conjugation and mass reversal. The mass generacy of neutrinos ν and ν' will strongly suggest the existence of mass parity conjugation operator C_p , which is expected to have the same character as that introduced by Lee and Yang.¹⁴⁾ Now let us assume the existence of C_p which obeys the conditions

$$[C_p, M]_- = 0 \quad (21)$$

and

$$[C_p, H_{int}]_- = 0 \quad (22)$$

where H_{int} denotes the interaction Hamiltonians. As is easily known, the condition (22) is not satisfied unless we take the suitably chosen interactions as H_{int} in (22). Therefore we understand the condition (22) to select the interactions H_{int} . Then the following Fermi interactions corresponding to $\mu \rightarrow e + \bar{\nu}$ and $\Lambda \rightarrow N + \bar{\nu} + \nu'$ only should be taken as H_{int} :

$$G(\bar{\psi}_e O_\alpha \psi_\mu) (\bar{\psi}_\nu O_\alpha \psi_\nu) + G(\bar{\psi}_e O_\alpha \psi_\mu) (\bar{\psi}_{\nu'} O_\alpha \psi_{\nu'})$$

and

$$G(\bar{\psi}_N O_\alpha \psi_{\Lambda(\Sigma)}) (\bar{\psi}_\nu O_\alpha \psi_{\nu'}) + G(\bar{\psi}_N O_\alpha \psi_{\Lambda(\Sigma)}) (\bar{\psi}_{\nu'} O_\alpha \psi_\nu) \text{ with } O_\alpha = \gamma_\mu (1 \pm \gamma_5)$$

The latter interactions are not invariant under MR(1) for ν' or ν when the leptons are C or D type fields under mass reversal, which means that the condition (21) is not satisfied. Thus the decays $K \rightarrow (\pi) + \nu + \nu'$ are impossible because they are supposed to proceed through these interactions.

The other unwanted but unforbidden processes will compete with the observed modes and will be shown to be hardly observed. Thus we can say without further discussions that Sachs' assumption that all the lepton processes must involve at least one neutrino is essentially proved.

Though the order of the four fermions in the interactions is fixed so far, we can allow it to change and impose the same relationships as before. In this case, more stringent restrictions are obtained. The answer is that only $V-A$ couplings¹⁵⁾, $S+P-T$ couplings and $S+P+(T/3)$ couplings are possible when the universal interactions (18) are taken into account. We can give its proof for β -decay assuming the original interaction to be of type (18), denoted shortly by $S_f^{(0)} (\psi_p^\dagger \psi_N; \phi_e^\dagger \phi_\nu)$. When we rearrange

the order to be $S_{\lambda}'^{(0)} = S_{\lambda}^{(0)}(\Psi_p, \phi_\nu; \phi_e, \Psi_N)$, this $S_{\lambda}'^{(0)}$ is shown to be related to $S_f^{(0)}$,¹⁹ as follows :

$$S_{\lambda}'^{(0)} = \sum_f a_{\lambda f} S_f^{(0)}, \quad S_f^{(0)}(\Psi_p^+ \Psi_N; \phi_e^+ \phi_\nu) = (\Psi_p \gamma_f \Psi_N) (\phi_e^+ (1 \pm \gamma_5) \gamma_f \phi_\nu) \quad (23)$$

where

$$(a_{\lambda f}) = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 1 & -\frac{1}{2} & 0 & \frac{1}{2} & -1 \\ \frac{3}{2} & 0 & -\frac{1}{2} & 0 & \frac{3}{2} \\ 1 & \frac{1}{2} & 0 & -\frac{1}{2} & -1 \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \quad (24)$$

If we apply MR(2) to both interactions, it is concluded that the (STP) coupling and the $(V-A)$, $S+P-T$ coupling in S_{λ}' go into $-(V-A)$, $-(S+P-T)$ in S_f and the $S+P+(T/3)$ couplings goes into $S+P-(T/3)$ in S_f in C-number theory. If it is required that the interchange mentioned above gives the minus sign to C-number universal interactions or matrix element, the decay mode $\mu \rightarrow e + e^+ + e^-$ is forbidden because this has the plus sign by the interchange. This is the other argument to rule out the mode $\mu \rightarrow e + e^+ + e^-$.

§ 5. Concluding remarks

The advantage of the theory of mass reversal over the Nishijima theory is that the various points are explained in a unified manner in the theory of mass reversal. For example, we need not assume the Lee-Yang neutrino to derive Sachs' assumption that the lepton processes must involve at least one neutrino. And we have more freedom to determine the coupling forms of Fermi interactions than in Nishijima's theory. The parity nonconservation does not always mean in our theory that the observed neutrino should be described by a two component Dirac field because the interactions such as (18) are allowed. From the standpoint of MR(1), Fujii and Iwata²⁰ have examined the polarization problem of the β -decay and μ - e decay and they found that there exists a slight difference in the case of β -decay between the two-component neutrino theory and ours. And it would be detect this difference when more precise experiments are done, although both theories so far explain the experiments well.

As is easily understood, the special mass reversal (MR(1) can be applied to Sakata's composite model for the baryons and mesons, and we can find that the space parity is conserved for the strong interactions, but it is violated for the weak interactions. In other words, all the Fermi interactions (including leptons) having only an even number of particles with same masses in themselves are strong and the others are weak. One remarkable fact is that in the application of MR(1) the Yukawa-type strong interactions (10) and the corresponding Sakata theory give different results. This fact shows, therefore, that there is a great difference between the Fermi interactions and the Yukawa ones, and that the mass reversal MR(1) will be the most fundamental operator among the various mass

reversals as far as the Fermi interactions are concerned, though the appropriate combination of the four fermions is required for the strong interactions. But the application of MR(1) was limited to the weak interactions in this paper because the charge independence symmetry is working on the strong interactions so effectively that no room for the additional introduction of MR(1) is likely to be found. As pointed out by Sakurai, the mass reversal MR(1) seems to be possible due to the fact that fermions have two components when the strong interactions are switched off. But this paper suggests that such an interpretation for MR(1) will be true only in a good approximation. The problem remaining is to look for the true meaning of MR(1).

Finally it will be added that the mass parity doublet for the neutrinos may imply that there are two kinds of the universal $V-A$ interactions with the slightly different coupling constants, one of which accompanies the neutrino ν and the other accompanies the neutrino ν' . There is the similar situation in the strong interactions that the K meson seems to interact with the baryons a little weaker than the pion does, which may be related with the fact that the mass parity of the K meson is opposite to that of the pion.

Acknowledgement

The author wishes to express his sincere thanks to Professor John S. Toll for the kind hospitality extended to him at the University of Maryland, and also to Drs. K. Senba and M. Yonezawa for helpful discussions and criticism on the manuscript. This work was supported by the Japan Ministry of Education.

References

- 1) T. Ouchi, K. Senba and M. Yonezawa, *Prog. Theor. Phys.* **15** (1956), 431.
- 2) K. Nishijima, *Nuovo Cimento* **5** (1957), 1349.
- 3) E. J. Konopinski and H.M. Mahmoud, *Phys. Rev.* **92** (1953), 1045.
- 4) K. Huang and F. E. Low, *Phys. Rev.* **109** (1958), 1400.
- 5) M. Goldhaber, L. Gordzins, and A. W. Sunyar, *Phys. Rev.* **109** (1958), 1015.
- 6) S. Hori and A. Wakasa, *Nuovo Cimento* **6** (1957), 304.
J. J. Sakurai, *Nuovo Cimento* **7** (1958), 649.
- 7) R. P. Feynman and M. Gell-Mann, *Phys. Rev.* **109** (1958), 193.
- 8) E. C. G. Sudarshan and R. E. Marshak, *Phys. Rev.* **109** (1958), 1860.
- 9) J. Tiomno, *Nuovo Cimento* **1**(1955), 226.
- 10) T. Ouchi, *Prog. Theor. Phys.* **17** (1957), 743. The definition of A -, B -, C -, and D - type. Field under mass reversal is given in this paper.
- 11) It is the essential point for the following discussions that g does not contain the explicit mass and therefore is a natural constant but, G need not be a constant.
- 12) R. G. Sachs, *Phys. Rev.* **99** (1955), 1573.
- 13) C. N. Yang and J. Tiomno, *Phys. Rev.* **79** (1950), 495.
- 14) The concept of parity doublet was first introduced by T. D. Lee and C. N. Yang (*Phys. Rev.* **102** (1956), 290). The mass parity doublet is the analogous to this.
- 15) K. Nishijima, *Phys. Rev.* **108** (1957), 907.
- 16) S. Sakata, *Prog. Theor. Phys.* **16** (1956), 686.
- 17) R. E. Marshak and E. C. G. Sudarshan, *Nuovo Cimento* **6** (1957), 1335.
- 18) More systematic argument was given by R. E. Behrends (*Phys. Rev.* **109** (1958), 2217.)
- 19) K. M. Case, *Phys. Rev.* **97** (1955), 810.
- 20) K. Fujii and K. Iwata, *Prog. Theor. Phys.* **13** (1957), 666.

Theory of Classical Fluids: Hyper-Netted Chain Approximation, I



— Formulation for a One-Component System —

Tohru MORITA

Physics Department, Tokyo Institute of Technology, Tokyo

(Received August 22, 1958)

Several attempts have been presented to take account of terms corresponding to large graphs in the cluster expansion formulae for the free energy and the radial distribution function, for the purpose of applying them to a gas of not small density or to a liquid. In this paper is proposed an approximation, which includes the approximations in the past as the first or second approximation, by taking account of graphs which can be easily evaluated by means of the Fourier transformation.

The graphs taken into account in our approximation are all those which can be constructed starting from a line $\bullet\text{---}\bullet$ or a ring  by a sequence of the processes to replace a constituent line $\bullet\text{---}\bullet$ by a watermelon ; in the course of the replacement, a line in a watermelon is also allowed to be replaced by a watermelon. The approximation will be called the hyper-netted chain approximation.

The formulae obtained in this approximation for the free energy and the radial distribution function are given by eqs. (29')–(30') and eq. (35), which demand the solution of a recurrence equation containing the Fourier transformations.

The expansion formulae for the free energy and the radial distribution function by means of the "hyper-netted chains" are also presented. They contain the results in the hyper-netted chain approximation as their leading terms.

Another set of the formulae in the hyper-netted chain approximation is given and compared with the theories for ionic systems in the past.

Applications to practical problems will be given in forthcoming papers.

§ 1. Introduction

We consider a uniform fluid of N identical particles in a volume V and temperature T ; the potential of the system is assumed to be the sum of potential $\phi(r)$ between every pair of particles.

The free energy A of the system is expanded in the virial series as¹⁾

$$\frac{A}{NkT} = \ln \left(\frac{h^2}{2\pi mkT} \right)^{3/2} \frac{\rho}{e} - \sum_{n=1}^{\infty} \frac{1}{n+1} \beta_n \rho^n \quad (1)$$

where m is the mass of a particle, h Planck's constant, k Boltzmann constant, $\rho = N/V$ is the density and β_n the irreducible integral given by

$$\beta_n = -\frac{1}{n!V} \iint \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_{n+1} \sum_{n+1 \geq i > j \geq 1} \prod f_{ij}^{(0)} \quad (2)$$

Sum over all products which are more than singly connected.

where

$$f_{ij}^{(0)} = f^{(0)}(r_{ij}) = \exp[-\phi(r_{ij})/kT] - 1. \quad (3)$$

As Mayer did,¹⁾ we express a term in the sum of the integrand of (2) by a graph consisted of numbered points, 1, 2, ..., $n+1$, representing the number of particles, 1, 2, ..., $n+1$, and of lines connecting points i and j representing the factors $f_{ij}^{(0)}$ in the summand.

To obtain a physically reasonable results by the help of this expansion for the case of a fluid of not small density, it will be inevitable to evaluate β_n for large n suitably. For this purpose, some attempts to take some suitable terms into account have been presented. Montroll and Mayer²⁾ presented the ring approximation, in which the graphs of the form of a ring are considered besides the graph of a line: the others neglected. It is of course suspected whether in this approximation β_n for each n can be evaluated correctly even in the qualitative sense. It seems, however, to give correct physical results, for this approximation is known to be derived also from another more physically intuitive approximation valid also for a fluid of not small density.³⁾ When we start from the virial expansion formula, in order to improve the approximation we have only to take more graphs into account. On the one hand, the netted chain or ring approximation was presented,⁴⁾ which takes account of the graphs which are obtained from a ring by replacing some of constituent line $\cdots \bullet \cdots \bullet \cdots$ by the figure $\cdots \bullet \triangle \bullet \cdots$, in addition. On the other hand, Abe⁵⁾ proposed the watermelon approximation, which takes into account all the graphs of chains which have two common end points, as shown in Fig. 1d, in addition to those taken into account by Montroll and Mayer.²⁾

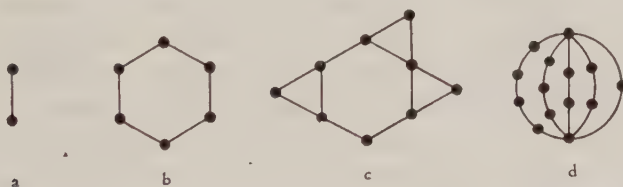


Fig. 1. a: second virial, b: ring, c: netted ring and d: watermelon approximation.

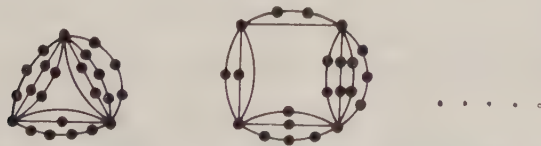


Fig. 2. Rings of watermelon.

Fig. 1 shows the typical graphs which have been taken into account in the approximations mentioned above. Abe has suggested that the approximation will be improved by taking account, in addition, of the graphs of ring of watermelons, as shown in Fig. 2. It is to be noted that this approximation will take into account far more graphs than

those considered in the netted ring approximation.

Thus, by generalizing the suggestion of Abe, we have a procedure to improve the approximation in a successive manner: We start from a line $\bullet\text{---}\bullet$ as the zeroth order approximation, construct rings of line, construct watermelons by adding to it sequences of line $\bullet\text{---}\bullet\text{---}\bullet$ between two points of a ring, and then construct rings of watermelon, and analogously we have watermelons of watermelon, rings of watermelon of watermelon,..... The purpose of the next section of this paper is to show that all these graphs can be taken into account for the formula of the free energy: this approximation will be called the "hyper-netted chain approximation". Several simple graphs considered in our approximation are shown in Fig. 3 and those not considered in Fig. 4.

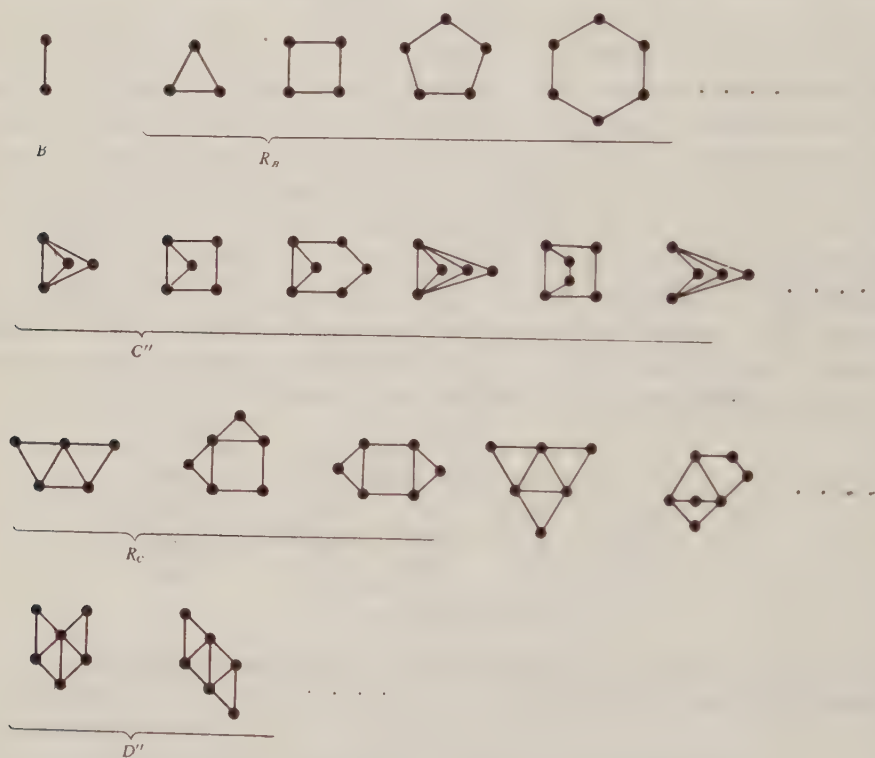


Fig. 3. Graphs considered in the hyper-netted chain approximation.



Fig. 4. Graphs not considered in the hyper-netted chain approximation.

In the above we have considered the approximation for the case of the free energy. In § 2, the formula for the free energy in our approximation will be derived, and in § 3

the formula for the radial distribution function in the same approximation will be obtained. In § 5, we will discuss the possible methods to obtain analogous but somewhat different formulae for the free energy and the radial distribution function in the same approximation and compare with the approximations presented in the past for ionic systems. In § 4, we insert the formulae for the free energy and the radial distribution function, without approximation, in which the contributions of clusters are summed in the spirit of the hyper-netted chain approximation.

§ 2. Free energy

For the purpose of ordering the graphs we introduce the process of "identification". Now, let us call a point a "junction" when three or more lines meet at the point, following Abe.⁵⁾ Let us call two lines to be "identifiable" when they have two common end points and both have no junction midway: For instance, in Fig. 5, two lines of *a* and four lines of *b* are identifiable, in *c* no identifiable lines are drawn, and in *d* upper two lines only are identifiable. Then, we will call the process of replacing identifiable lines by a line the "identification".

By means of this terminology, the graphs considered in our approximation are all those which can be reduced to a line $\bullet\text{---}\bullet$ or a ring \bigcirc by a sequence of identifications.* The other graphs will be considered in § 4.

We group the graphs to be considered by the times of identifications needed to reduce them to a line or a ring. The times of identification are measured by identifying all the identifiable lines in a graph at a time; so that the graphs of C'' in Fig. 3 are reduced to a line by one time of identification, those of R_c to a ring by one time of identification and those of D'' to a line by two times of identification. We calculate the contribution to the free energy

$$-\frac{A'}{NkT} = \sum_{n=1}^{\infty} \frac{1}{n+1} \beta_n \rho^n = \sum_{n=1}^{\infty} \frac{\rho^n}{(n+1)!} \int \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_{n+1} \sum_{\substack{n+1 \geq i > j \geq 1}} \prod f_{ij}^{(0)} \quad (4)$$

Sum over all products which are more than singly connected.

for each group separately, and add them up at last.

We will start with graphs for which the times of identification are small and proceed to those with larger and larger times of identification. First, we consider the graph which is a line at the zeroth time of identification. A line $\bullet\text{---}\bullet$, representing $f_{ij}^{(0)}$, will be denoted as *B* in the graphs in the following. The contribution of *B* to (4) is

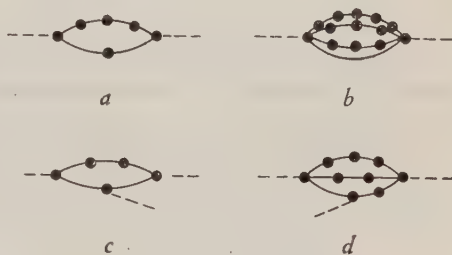


Fig. 5.

Here, - - - indicates a line or lines connected to another parts of the graph.

* As a ring can be reduced further to a line by another identification, these may be said to be all graphs which can be reduced to a line by a sequence of identifications.

$$\left(-\frac{A'}{NkT}\right)_B = -\frac{\rho}{2V} \iint d\mathbf{r}_1 d\mathbf{r}_2 f^{(0)}(r_{12}) = \frac{\rho}{2} F^{(0)}(0), \tag{5}$$

where $F^{(0)}(\mathbf{k})$ is the Fourier transform of $f^{(0)}(\mathbf{r})$: in general, we put

$$F^{(n)}(\mathbf{k}) = \int d\mathbf{r} e^{i\mathbf{k}\cdot\mathbf{r}} f^{(n)}(\mathbf{r}) \tag{6}$$

for $f^{(0)}(\mathbf{r})$ defined by (3) and $f^{(n)}(\mathbf{r})$ with $n \geq 1$ to be defined in the following.

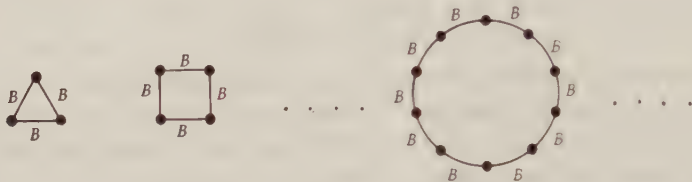


Fig. 6. R_B -rings of B .

Next, we consider the graphs of ring, the total of which, shown in Fig. 6, will be called R_B -rings of B . The contribution of R_B to (4) is

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{R_B} &= \sum_{n=2}^{\infty} \frac{\rho^n n!}{2(n+1)! V} \iint \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_{n+1} f^{(0)}(r_{12}) f^{(0)}(r_{23}) \cdots f^{(0)}(r_{n+1,1}) \\ &= \frac{1}{2\rho V} \sum_{\mathbf{k}} \sum_{n=3}^{\infty} \frac{1}{n} \rho^n F^{(0)}(\mathbf{k})^n \\ &= \frac{1}{2\rho V} \sum_{\mathbf{k}} \left\{ -\ln(1 - \rho F^{(0)}(\mathbf{k})) - \rho F^{(0)}(\mathbf{k}) - \frac{1}{2} \rho^2 F^{(0)}(\mathbf{k})^2 \right\}; \end{aligned} \tag{7}$$

as has been evaluated by Montroll and Mayer.

Next, we consider the graphs which become a line by one time of identification. Now let us call each of the sequences of B shown in Fig. 7 B_S , and define the groups

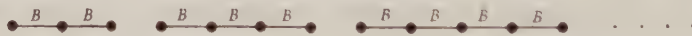


Fig. 7. B_S -sequences of B .

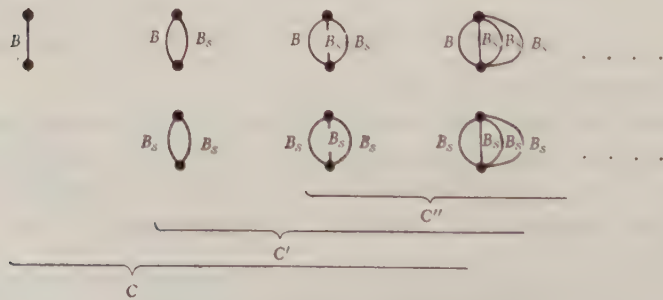


Fig. 8. C, C', C'' -watermelons of B .

of graphs C , C' and C'' as in Fig. 8; B appears besides B_s because one and only one B is allowed to exist between a pair of points. The contribution of C'' to (4) is

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{C''} &= \frac{1}{2} \sum_{\mathbf{k}_1} \sum_{\substack{\mathbf{k}_2 \\ \Sigma \mathbf{k}_i = 0}} \cdots \left\{ \rho F^{(0)}(\mathbf{k}_1) \left[\frac{1}{2!V} H^{(0)}(\mathbf{k}_2) H^{(0)}(\mathbf{k}_3) \right. \right. \\ &\quad \left. \left. + \frac{1}{3!V^2} H^{(0)}(\mathbf{k}_2) H^{(0)}(\mathbf{k}_3) H^{(0)}(\mathbf{k}_4) + \cdots \right] \right. \\ &\quad \left. + \rho \left[\frac{1}{3!V^2} H^{(0)}(\mathbf{k}_1) H^{(0)}(\mathbf{k}_2) H^{(0)}(\mathbf{k}_3) \right. \right. \\ &\quad \left. \left. + \frac{1}{4!V^3} H^{(0)}(\mathbf{k}_1) H^{(0)}(\mathbf{k}_2) H^{(0)}(\mathbf{k}_3) H^{(0)}(\mathbf{k}_4) + \cdots \right] \right\} \quad (8) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2} \int d\mathbf{r} \left\{ \rho f^{(0)}(\mathbf{r}) \left[\frac{1}{2!} h^{(0)}(\mathbf{r})^2 + \frac{1}{3!} h^{(0)}(\mathbf{r})^3 + \cdots \right] \right. \\ &\quad \left. + \rho \left[\frac{1}{3!} h^{(0)}(\mathbf{r})^3 + \frac{1}{4!} h^{(0)}(\mathbf{r})^4 + \cdots \right] \right\} \\ &= \frac{\rho}{2} \int d\mathbf{r} \left\{ f^{(0)}(\mathbf{r}) [e^{h^{(0)}(\mathbf{r})} - 1 - h^{(0)}(\mathbf{r})] \right. \\ &\quad \left. + e^{h^{(0)}(\mathbf{r})} - 1 - h^{(0)}(\mathbf{r}) - \frac{1}{2!} h^{(0)}(\mathbf{r})^2 \right\} \quad (8') \end{aligned}$$

where

$$H^{(0)}(\mathbf{k}) = \rho F^{(0)}(\mathbf{k})^2 + \rho^2 F^{(0)}(\mathbf{k})^3 + \cdots = \frac{\rho F^{(0)}(\mathbf{k})^2}{1 - \rho F^{(0)}(\mathbf{k})} \quad (9)$$

and its Fourier transform

$$h^{(0)}(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{r}} H^{(0)}(\mathbf{k}) \quad (10)$$

represents the total of the propagation lines of B_s^* ; as has been obtained by Abe.⁽⁷⁾ This calculation has been performed by considering that: 1) The power of ρ is equal to the number of points in the graph minus one. 2) We may transform the integration by the coordinate \mathbf{x} of the upper end point of a graph of C'' as

$$\begin{aligned} &\int d\mathbf{x} f_1(\mathbf{x}, \cdots) f_2(\mathbf{x}, \cdots) \cdots f_n(\mathbf{x}, \cdots) \\ &= \int d\mathbf{x} \int d\mathbf{x}_1 \int d\mathbf{x}_2 \cdots \int d\mathbf{x}_n \frac{1}{V^n} \sum_{\mathbf{k}_1} \sum_{\mathbf{k}_2} \cdots \sum_{\mathbf{k}_n} \\ &\quad e^{i\mathbf{k}_1 \cdot (\mathbf{x} - \mathbf{x}_1)} e^{i\mathbf{k}_2 \cdot (\mathbf{x} - \mathbf{x}_2)} \cdots e^{i\mathbf{k}_n \cdot (\mathbf{x} - \mathbf{x}_n)} f_1(\mathbf{x}_1, \cdots) f_2(\mathbf{x}_2, \cdots) \cdots f_n(\mathbf{x}_n, \cdots) \end{aligned}$$

* By "propagation line", we mean a graph which consists of lines, watermelons, ..., with two end points and is connected with another parts of the graph by these two end points alone, and by "propagation", we imply the factor by which it contributes to (4),

$$= \frac{1}{V^{n+1}} \sum_{k_1} \sum_{k_2} \dots \sum_{k_n} e^{ik_1 \cdot x_1} e^{ik_2 \cdot x_2} \dots e^{ik_n \cdot x_n} f_1(x_1, \dots) f_2(x_2, \dots) \dots f_n(x_n, \dots), \tag{11}$$

then we can integrate over the coordinates of the midway points by a convolution as (8).
3) Factors $1/2$ and $1/2!, 1/3!, \dots$ are needed because of the fact that, if all possible graphs in Fig. 7 are inserted in B 's in C'' and the numbers of particles are permuted in $(n+1)!$ ways, then $2 \cdot 2!, 2 \cdot 3!, \dots$ identical graphs appear. The details of this calculation are given in Abe's paper.

Next, we consider the graphs which reduce to a ring by one time of identification. These graphs, to be referred to as R_c -rings of C , are the total of those graphs in Fig. 9



Fig. 9. R_c^* -rings of C .

for which at least two edges are occupied by a propagation of C' . That is, R_c is R_c^* , the total of Fig. 9, minus the graphs for which all C 's are equal to B and the graphs for which one of C is C' and all others are B . At first we calculate the contribution to (4) of R_c^*

$$\left(-\frac{A'}{NkT}\right)_{R_c^*} = \frac{1}{2\rho V} \sum_k \sum_{n=3}^\infty \frac{1}{n} \rho^n F^{(1)}(k)^n \tag{12}$$

$$= \frac{1}{2\rho V} \sum_k \left\{ -\ln(1 - \rho F^{(1)}(k)) - \rho F^{(1)}(k) - \frac{1}{2} \rho^2 F^{(1)}(k)^2 \right\} \tag{12'}$$

where $f^{(1)}(\mathbf{r})$ and its Fourier transform $F^{(1)}(\mathbf{k})$ represent the propagation by the whole graphs of C in Fig. 8 :

$$f^{(1)}(\mathbf{r}) = f^{(0)}(\mathbf{r}) \sum_{n=0}^\infty \frac{1}{n!} h^{(0)}(\mathbf{r})^n + \sum_{n=2}^\infty \frac{1}{n!} h^{(0)}(\mathbf{r})^n \tag{13}$$

$$= f^{(0)}(\mathbf{r}) e^{h^{(0)}(\mathbf{r})} + e^{h^{(0)}(\mathbf{r})} - 1 - h^{(0)}(\mathbf{r}). \tag{13'}$$

To obtain eq. (12) with (13), we have taken the following procedure. First, we consider the contribution of a graph of the type just considering. We integrate first over all the coordinates except two end points for every watermelon constituting the ring. Then we have a term in the summand of (13) for each edge or watermelon of the ring, having the distance between the two end points of it as \mathbf{r} . Then, we can easily integrate over the coordinates of vertices of the ring by an aid of the convolution. We have then only to sum the contributions of all possible graphs just in consideration, where cares must be taken not to add the contribution of a graph many times. The coefficient can be evaluated

as follows. First, we construct all possible graphs by inserting one of C 's in each edge of the ring of C in Fig. 9 independently of the occupation of the other edges and, in turn, one of B_s in each line of C independently of the occupation of the other lines; and next we take all possible permutation of the numeration of particles in the graphs, then we have (number of particles)! terms from the graph above constructed. There will appear many graphs indistinguishable with each other not only in the form but also in the numbers of molecules. To obtain the contribution to (4), we take this into account by introducing the factor $1/2n$ in (12) for the rotational and reflective symmetries of the ring and the factors $1/2!$, $1/3!$, ... in (13) for the invariance by the interchange of 2, 3, ... lines of B_s in the watermelon C of the edges.

In eq. (12) we divide $F^{(1)}(\mathbf{k})$ into two parts due to B and C' as

$$F^{(1)}(\mathbf{k}) = F^{(0)}(\mathbf{k}) + \Delta F^{(1)}(\mathbf{k}), \quad (14)$$

then the term of the zeroth order in $\Delta F^{(1)}(\mathbf{k})$ represents the contribution to $(-A'/NkT)_{RC^*}$ from the graphs for which C 's are all equal to B and the term of the first order in $\Delta F^{(1)}(\mathbf{k})$ from the graphs for which one C is C' and all others are B . The former is equal to $(-A'/NkT)_{RB}$ (cf. eq. (7)) and the latter is obtained as

$$\begin{aligned} \frac{1}{2\rho V} \sum_{\mathbf{k}} \sum_{n=1}^{\infty} \rho^n F^{(0)}(\mathbf{k})^{n-1} \Delta F^{(1)}(\mathbf{k}) &= -\frac{\rho}{2V} \sum_{\mathbf{k}} H^{(0)}(\mathbf{k}) \Delta F^{(1)}(\mathbf{k}) \\ &= -\frac{\rho}{2V} \sum_{\mathbf{k}} H^{(0)}(\mathbf{k}) [F^{(1)}(\mathbf{k}) - F^{(0)}(\mathbf{k})] = \frac{\rho}{2} \int d\mathbf{r} h^{(0)}(\mathbf{r}) [f^{(1)}(\mathbf{r}) - f^{(0)}(\mathbf{r})], \end{aligned} \quad (15)$$

where the relation $h^{(0)}(\mathbf{r}) = h^{(0)}(-\mathbf{r})$ has been used. As the result we have

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{RC'} &= \left(-\frac{A'}{NkT}\right)_{RC^*} - \left(-\frac{A'}{NkT}\right)_{RB} \\ &\quad - \frac{\rho}{2} \int d\mathbf{r} h^{(0)}(\mathbf{r}) [f^{(1)}(\mathbf{r}) - f^{(0)}(\mathbf{r})]. \end{aligned} \quad (16)$$

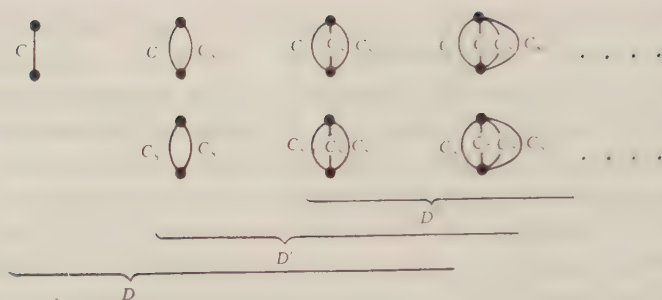
With the aid of eq. (13), (8') can be rewritten as

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{C'} &= -\frac{\rho}{2} [F^{(1)}(0) - F^{(0)}(0)] - \frac{\rho}{2} \int d\mathbf{r} f^{(0)}(\mathbf{r}) h^{(0)}(\mathbf{r}) \\ &\quad - \frac{\rho}{4} \int d\mathbf{r} h^{(0)}(\mathbf{r})^2. \end{aligned} \quad (8'')$$

Next, we proceed to the graphs which become a line by two times of identification. Now we construct the sequence C_s^* of C as in Fig. 10. Then we define C_s by $C_s = C_s^* - B_s$. Then we construct the watermelons of C and C_s as in Fig. 11. The graphs of D'' are the graphs to be considered here. The contribution to (4) is obtained as



Fig. 10. C_s^* —sequences of C . $C_s = C_s^* - B_s$.

Fig. 11. D , D' , D'' —watermelons of C .

$$\left(-\frac{A'}{NkT}\right)_{D''} = \frac{\rho}{2} \int d\mathbf{r} \left\{ f^{(1)}(\mathbf{r}) [e^{\Delta h^{(1)}(\mathbf{r})} - 1 - \Delta h^{(1)}(\mathbf{r})] \right. \\ \left. + e^{\Delta h^{(1)}(\mathbf{r})} - 1 - \Delta h^{(1)}(\mathbf{r}) - \frac{1}{2!} [\Delta h^{(1)}(\mathbf{r})]^2 \right\}, \quad (17)$$

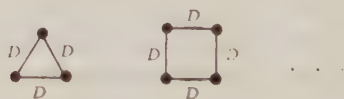
just like when the contribution of C'' was obtained, where $\Delta h^{(1)}(\mathbf{r})$ and its Fourier transform of $\Delta H^{(1)}(\mathbf{k})$, defined by

$$\Delta H^{(1)}(\mathbf{k}) = H^{(1)}(\mathbf{k}) - H^{(0)}(\mathbf{k}) \quad (18)$$

are the propagation by C_S , $H^{(1)}(\mathbf{k})$ being the propagation by C_S^* :

$$H^{(1)}(\mathbf{k}) = \frac{\rho F^{(1)}(\mathbf{k})^2}{1 - \rho F^{(1)}(\mathbf{k})}. \quad (19)$$

Next, we calculate the contributions of R_D , the rings of D with at least two D 's equal to D' . Just like for the case of R_C , we calculate it by subtracting from

Fig. 12. R_D^* —rings of D .

$$\left(-\frac{A'}{NkT}\right)_{R_D^*} = \frac{1}{2\rho V} \sum_{\mathbf{k}} \sum_{n=2}^{\infty} \frac{1}{n} \rho^n F^{(2)}(\mathbf{k})^n \\ = \frac{1}{2\rho V} \sum_{\mathbf{k}} \left\{ -\ln(1 - \rho F^{(2)}(\mathbf{k})) - \rho F^{(2)}(\mathbf{k}) - \frac{1}{2} \rho^2 F^{(2)}(\mathbf{k})^2 \right\}, \quad (20)$$

where $F^{(2)}(\mathbf{k})$ and $f^{(2)}(\mathbf{r})$ represent the propagation by D :

$$f^{(2)}(\mathbf{r}) = f^{(1)}(\mathbf{r}) e^{\Delta h^{(1)}(\mathbf{r})} + e^{\Delta h^{(1)}(\mathbf{r})} - 1 - \Delta h^{(1)}(\mathbf{r}), \quad (21)$$

the terms of the zeroth order of $F^{(2)}(\mathbf{k}) - F^{(1)}(\mathbf{k})$, equal to $(-A'/NkT)_{R_C^*}$, and the term of the first order, equal to

$$-\frac{\rho}{2V} \sum_k H^{(1)}(\mathbf{k}) [F^{(2)}(\mathbf{k}) - F^{(1)}(\mathbf{k})] = \frac{\rho}{2} \int d\mathbf{r} h^{(1)}(\mathbf{r}) [f^{(2)}(\mathbf{r}) - f^{(1)}(\mathbf{r})]. \quad (22)$$

As the result, we have

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{ED} &= \left(-\frac{A'}{NkT}\right)_{ED^*} - \left(-\frac{A'}{NkT}\right)_{BC^*} \\ &\quad - \frac{\rho}{2} \int d\mathbf{r} h^{(1)}(\mathbf{r}) [f^{(2)}(\mathbf{r}) - f^{(1)}(\mathbf{r})]. \end{aligned} \quad (23)$$

The contribution of D'' to (4) can be rewritten as, by substituting (21) in (17),

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{D''} &= \frac{\rho}{2} [F^{(2)}(0) - F^{(1)}(0)] \\ &\quad - \frac{\rho}{2} \int d\mathbf{r} f^{(1)}(\mathbf{r}) \Delta h^{(1)}(\mathbf{r}) - \frac{\rho}{4} \int d\mathbf{r} [\Delta h^{(1)}(\mathbf{r})]^2, \end{aligned} \quad (17')$$

and so forth.

In this way, the contribution of the graphs $B^{(n)} \text{---} B^{(0)} = B$, $B^{(1)} = C''$, $B^{(2)} = D''$, ... —which are reduced to a line by n times of identification is obtained as

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{B^{(n)}} &= \frac{\rho}{2} [F^{(n)}(0) - F^{(n-1)}(0)] \\ &\quad - \frac{\rho}{2} \int d\mathbf{r} f^{(n-1)}(\mathbf{r}) \Delta h^{(n-1)}(\mathbf{r}) - \frac{\rho}{4} \int d\mathbf{r} [\Delta h^{(n-1)}(\mathbf{r})]^2 \end{aligned} \quad (24)$$

where $F^{(-1)}(\mathbf{k}) \equiv 0$, $\Delta h^{(-1)}(\mathbf{r}) \equiv h^{(-1)}(\mathbf{r}) \equiv 0$. The contribution of the graph $R^{(n)} \text{---} R^{(0)} = R_R$, $R^{(1)} = R_C$, $R^{(2)} = R_D$, ... —which reduce to a ring by n times of identification, is

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{R^{(n)}} &= \left(-\frac{A'}{NkT}\right)_{R^{(n)}^*} - \left(-\frac{A'}{NkT}\right)_{R^{(n-1)}^*} \\ &\quad - \frac{\rho}{2} \int d\mathbf{r} h^{(n-1)}(\mathbf{r}) [f^{(n)}(\mathbf{r}) - f^{(n-1)}(\mathbf{r})] \end{aligned} \quad (25)$$

with

$$\left(-\frac{A'}{NkT}\right)_{R^{(n)}^*} = \frac{1}{2\rho V} \sum_k \left\{ -\ln(1 - \rho F^{(n)}(\mathbf{k})) - \rho F^{(n)}(\mathbf{k}) - \frac{1}{2} \rho^2 F^{(n)}(\mathbf{k})^2 \right\}. \quad (26)$$

Summing these up to $B^{(n)}$ and $R^{(n)}$, we have

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{\sum_{s=0}^n (B^{(s)} + R^{(s)})} &= \frac{\rho}{2} F^{(n)}(0) \\ &\quad + \frac{1}{2\rho V} \sum_k \left\{ -\ln(1 - \rho F^{(n)}(\mathbf{k})) - \rho F^{(n)}(\mathbf{k}) - \frac{1}{2} \rho^2 F^{(n)}(\mathbf{k})^2 \right\} \\ &\quad - \frac{\rho}{2} \int d\mathbf{r} f^{(n)}(\mathbf{r}) f^{(n-1)}(\mathbf{r}) - \frac{\rho}{4} \int d\mathbf{r} \sum_{s=0}^{n-1} [\Delta h^{(s)}(\mathbf{r})]^2. \end{aligned} \quad (27)$$

Summing these up to $R^{(n)}$ and $B^{(n+1)}$, we have

$$\begin{aligned} \left(-\frac{A'}{NkT} \right)_{\sum_{s=0}^n (B^{(s)} + R^{(s)}) + B^{(n+1)}} &= \frac{\rho}{2} F^{(n+1)}(0) \\ &+ \frac{1}{2\rho V} \sum_{\mathbf{k}} \left\{ -\ln(1 - \rho F^{(n)}(\mathbf{k})) - \rho F^{(n)}(\mathbf{k}) - \frac{1}{2} \rho^2 F^{(n)}(\mathbf{k})^2 \right\} \\ &- \frac{\rho}{2} \int d\mathbf{r} f^{(n)}(\mathbf{r}) h^{(n)}(\mathbf{r}) - \frac{\rho}{4} \int d\mathbf{r} \sum_{s=0}^n [Jh^{(s)}(\mathbf{r})]^2. \end{aligned} \quad (28)$$

In the limit of $n \rightarrow \infty$, these both tend to

$$\begin{aligned} \left(-\frac{A'}{NkT} \right)_{\text{HNC}} &= \frac{\rho}{2} \mathcal{F}(0) + \frac{1}{2\rho V} \sum_{\mathbf{k}} \left\{ -\ln(1 - \rho \mathcal{F}(\mathbf{k})) - \rho \mathcal{F}(\mathbf{k}) - \frac{1}{2} \rho^2 \mathcal{F}(\mathbf{k})^2 \right\} \\ &- \frac{\rho}{2} \int d\mathbf{r} \mathcal{f}(\mathbf{r}) \mathcal{h}(\mathbf{r}) - \frac{\rho}{4} \int d\mathbf{r} \sum_{s=0}^{\infty} [Jh^{(s)}(\mathbf{r})]^2 \end{aligned} \quad (29)$$

with

$$\begin{aligned} \mathcal{f}(\mathbf{r}) &= \lim_{n \rightarrow \infty} f^{(n)}(\mathbf{r}), \quad \mathcal{h}(\mathbf{r}) = \lim_{n \rightarrow \infty} h^{(n)}(\mathbf{r}), \quad \mathcal{F}(\mathbf{k}) = \lim_{n \rightarrow \infty} F^{(n)}(\mathbf{k}) \\ f^{(n+1)}(\mathbf{r}) &= f^{(n)}(\mathbf{r}) e^{\Delta h^{(n)}(\mathbf{r})} + e^{\Delta h^{(n)}(\mathbf{r})} - 1 - Jh^{(n)}(\mathbf{r}) \\ \Delta h^{(n)}(\mathbf{r}) &= h^{(n)}(\mathbf{r}) - h^{(n-1)}(\mathbf{r}) \\ h^{(n)}(\mathbf{r}) &= \frac{1}{V} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{r}} H^{(n)}(\mathbf{k}) \\ H^{(n)}(\mathbf{k}) &= \frac{\rho F^{(n)}(\mathbf{k})^2}{1 - \rho F^{(n)}(\mathbf{k})} \\ F^{(n)}(\mathbf{k}) &= \int d\mathbf{r} e^{i\mathbf{k} \cdot \mathbf{r}} f^{(n)}(\mathbf{r}) \\ f^{(0)}(\mathbf{r}) &= e^{-\phi(\mathbf{r})/kT} - 1, \quad h^{(-1)}(\mathbf{r}) = 0. \end{aligned} \quad (30)$$

When the potential $\phi(\mathbf{r})$ is a function of the distance $r = |\mathbf{r}|$ alone, then so are $f^{(n)}(\mathbf{r})$ and $h^{(n)}(\mathbf{r})$, and $F^{(n)}(\mathbf{k})$ and $H^{(n)}(\mathbf{k})$ are functions of $k = |\mathbf{k}|$ alone. Then the above set of formulae can be written as

$$\begin{aligned} \left(-\frac{A'}{NkT} \right)_{\text{HNC}} &= \frac{\rho}{2} \mathcal{F}(0) + \frac{1}{4\pi^2 \rho} \int_0^\infty k^2 dk \left\{ -\ln(1 - \rho \mathcal{F}(k)) - \rho \mathcal{F}(k) - \frac{1}{2} \rho^2 \mathcal{F}(k)^2 \right\} \\ &- 2\pi\rho \int_0^\infty r^2 dr \mathcal{f}(r) \mathcal{h}(r) - \pi\rho \int_0^\infty r^2 dr \sum_{s=0}^\infty [Jh^{(s)}(r)]^2 \end{aligned} \quad (29')$$

with

$$\begin{aligned}
f(r) &= \lim_{n \rightarrow \infty} f^{(n)}(r), \quad h(r) = \lim_{n \rightarrow \infty} h^{(n)}(r), \quad \mathcal{F}(k) = \lim_{n \rightarrow \infty} F^{(n)}(k) \\
f^{(n+1)}(r) &= f^{(n)}(r) e^{\Delta h^{(n)}(r)} + e^{\Delta h^{(n)}(r)} - 1 - \Delta h^{(n)}(r) \\
\Delta h^{(n)}(r) &= h^{(n)}(r) - h^{(n-1)}(r) \\
h^{(n)}(r) &= \frac{1}{4\pi^2 r} \int_{-\infty}^{\infty} k dk \sin kr H^{(n)}(k) \\
H^{(n)}(k) &= \frac{\rho F^{(n)}(k)^2}{1 - \rho F^{(n)}(k)} \\
F^{(n)}(k) &= \frac{2\pi}{k} \int_{-\infty}^{\infty} r dr \sin kr f^{(n)}(r) \\
f^{(0)}(r) &= f^{(0)}(-r) = e^{-\phi(r)/kT} - 1, \quad h^{(-1)}(r) \equiv 0.
\end{aligned} \tag{30'}$$

Thus to get the numerical results in our approximation we have only to repeat the Fourier transformations starting from a given $f^{(0)}(r) = \exp(-\phi(r)/kT) - 1$.

§ 3. Radial distribution function

In the previous section, we have developed the hyper-netted chain approximation for the case of the free energy, from which all thermodynamical quantities can be derived. While, when we know the radial distribution function $g(r)$ we can calculate the pressure p and the internal energy E directly by

$$\frac{p}{\rho kT} = 1 - \frac{\rho}{6kT} \int_0^{\infty} g(r) \frac{d\phi(r)}{dr} 4\pi r^3 dr, \tag{31}$$

$$\frac{E}{\frac{3}{2} NkT} = 1 + \frac{\rho}{3kT} \int_0^{\infty} g(r) \phi(r) 4\pi r^2 dr. \tag{32}$$

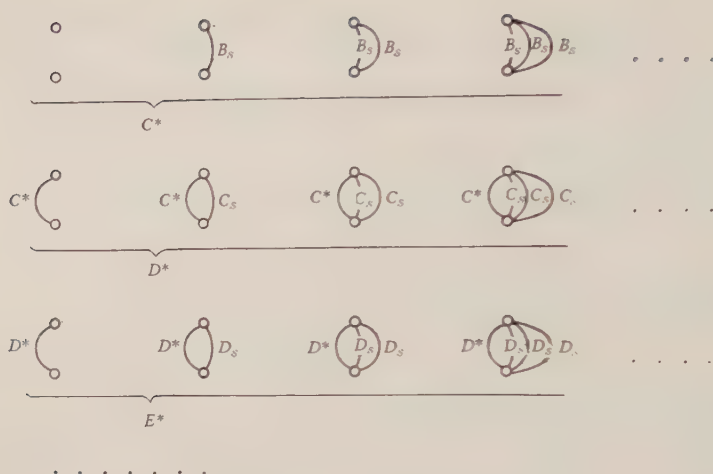
The virial expansion formula for the radial distribution function $g(r)$ is given by^{(2), (4)}

$$g(r_{12}) = e^{-\phi(r_{12})/kT} \left\{ 1 + \sum_{m=1}^{\infty} H_{2m}(\mathbf{r}_1, \mathbf{r}_2) \rho^m \right\} \tag{33}$$

$$H_{2m}(\mathbf{r}_1, \mathbf{r}_2) = (1/m!) \iint \cdots \int d\mathbf{r}_3 \cdots d\mathbf{r}_{m+2} \sum_{\substack{m+2 \geq i > j \geq 3 \\ m+2 \geq k \geq 1 \\ 2 \geq i, k \geq 1}} \prod f_{ij}^{(0)} f_{k\ell}^{(0)} \tag{34}$$

Sum over all connected products for which each molecule of the set $\{\mathbf{r}_3, \dots, \mathbf{r}_{m+2}\}$ is connected to \mathbf{r}_1 and \mathbf{r}_2 by an independent path.

In this case also, we can group the graphs by the times of identification needed to reduce them to a line $\text{---}\text{---}\text{---}$. In Fig. 13, C^* contains the graphs of one time and two times, D^* those up to three times, and so on. The graphs which cannot be reduced to

Fig. 13. C^* , D^* , E^* , ...

a line by identifications will be considered in § 4.

Taking the graphs C^* only into account in (34), we have*

$$[g(r)]_{C^*} = e^{-\phi(r)/kT + \bar{h}^{(0)}(r)}.$$

Taking the graphs D^* into account, we have

$$[g(r)]_{D^*} = e^{-\phi(r)/kT + \bar{h}^{(0)}(r) + \Delta \bar{h}^{(1)}(r)} = e^{-\phi(r)/kT + \bar{h}^{(1)}(r)}.$$

In the approximation of E^* , we have

$$[g(r)]_{E^*} = e^{-\phi(r)/kT + \bar{h}^{(2)}(r)},$$

and so forth.

As the result, we have at last

$$[g(r)]_{\text{HNC}} = \exp[-\phi(r)/kT + \bar{h}(r)]; \quad (35)$$

$\bar{h}(r)$ is the solution of (30').

That this formula for $g(r)$ is in the same approximation as the formula for the free energy in the previous section can be confirmed by taking in (2) only the graphs considered in the previous section when we get (34) from \mathcal{A} of (2) or (4) by the formula due to Bogolyubov⁷⁾ and Hiroike⁸⁾:

$$g(r) = \frac{2V}{N^2} \frac{\partial \mathcal{A}}{\partial \phi(r)}.$$

* After the completion of the manuscript of this paper, the author was noticed that Meeron^{13), 14)} had presented some papers which are related to the summation of the cluster integrals for the distribution functions and potentials of average force. His result in ref. 14) is identical to ours in this approximation C^* , with corrections in terms of $b^{(0)}$, $f^{(0)}$, and $f^{(1)}$ in our notation. His result for $g(r)$ in ref. 13) for ionic solutions is also in an analogous approximation.

If we follow the line of Meeron⁽¹¹⁾ and are to calculate the potential of average force by means of the formula

$$\ln g(r_{12}) = -\frac{\phi(r_{12})}{kT} + \sum_{m=1}^{\infty} K_{2m}(\mathbf{r}_1, \mathbf{r}_2) \rho^m \quad (33')$$

$$K_{2m}(\mathbf{r}_1, \mathbf{r}_2) = (1/m!) \int \cdots \int d\mathbf{r}_3 \cdots d\mathbf{r}_{m+2} \sum_{\substack{m+2 \geq i > j \geq 3 \\ m+2 \geq k \geq 3 \\ 2 \geq i, k \geq 1}} \Pi f_{ij}^{(0)} f_{k\kappa}^{(0)}, \quad (34')$$

Sum over all products for which each molecule of the set $\{\mathbf{r}_3, \dots, \mathbf{r}_{m+2}\}$ is connected to \mathbf{r}_1 and \mathbf{r}_2 by an independent path and also the molecules $\{\mathbf{r}_3, \dots, \mathbf{r}_{m+2}\}$ are connected to each other even when we erase the lines ending at \mathbf{r}_1 and \mathbf{r}_2 .

the graphs to be considered to derive eq. (35) are obviously the total of

$$\begin{array}{c} \circ \\ \circ \end{array} B_s, \quad \begin{array}{c} \circ \\ \circ \end{array} C_s, \quad \begin{array}{c} \circ \\ \circ \end{array} D_s, \quad \dots \dots \dots$$

§ 4. Expansion formulae for the free energy and the radial distribution function by means of the hyper-netted chains

In the preceding sections, we have considered the graphs which are reduced to a line or a ring by a sequence of identifications. In this section, we consider the remaining graphs. We group the latter by the graphs of junctions which are obtained from them after a sequence of identifications is applied until they come to have no identifiable lines and then the points other than junctions are neglected. Then we calculate the contributions of these groups to (4) and (33) with (34) along the line used in the previous sections and obtain for the formulae for the free energy and the radial distribution function, without approximation, as follows:

$$-\frac{A'}{NkT} = \left(-\frac{A'}{NkT} \right)_{\text{HNC, given by eq. (29')}} + \sum_{n=1}^{\infty} \frac{\rho^{n-1}}{n! V} \int \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_n \sum_{n \geq i > j \geq 1} \Pi (\mathcal{f}_{ij} + \mathcal{h}_{ij}), \quad (36)$$

Sum over all products which consists of junctions alone and are more than singly connected.

$$g(r_{12}) = [g(r_{12})]_{\text{HNC, given by eq. (35)}}$$

$$\times \left\{ 1 + \sum_{m=2}^{\infty} \frac{\rho^m}{m!} \int \cdots \int d\mathbf{r}_3 \cdots d\mathbf{r}_{m+2} \sum_{\substack{m+2 \geq i > j \geq 3 \\ m+2 \geq k \geq 3 \\ 2 \geq i, k \geq 1}} \Pi (\mathcal{f}_{ij} + \mathcal{h}_{ij}) (\mathcal{f}_{k\kappa} + \mathcal{h}_{k\kappa}) \right\}, \quad (37)$$

Sum over all products for which each molecule of the set $\{\mathbf{r}_3, \dots, \mathbf{r}_{m+2}\}$ is a junction and is connected to \mathbf{r}_1 and \mathbf{r}_2 by an independent path.

where

$$f_{ij}+h_{ij}=f(r_{ij})+h(r_{ij}) ;$$

$f(r)$ and $h(r)$ are the solutions of eqs. (30') and represent the propagations of the graphs, Z and Z_n , which are the limits of the series B, C, D, \cdots and B_n, C_n, D_n, \cdots , respectively. The graphs corresponding to the products of the correction terms of eq. (36) and of eq. (37) are shown in Fig. 14 and in Fig. 15, respectively, for small n .



Fig. 14. A line $\bullet\cdots\bullet$ represents the hyper-netted chain, $Z+Z_n$; the limit of the series, $B+B_n, C+C_n, \cdots$



Fig. 15. A line $\bullet\cdots\bullet$ represents the hyper-netted chain, $Z+Z_n$.

It is to be noted that eq. (37) can be written also as: (cf. eq. (33') with (34'))

$$\ln g(r_{12}) = -\frac{\phi(r_{12})}{kT} + h(r_{12}) + \sum_{m=2}^{\infty} \frac{(-1)^m}{m!} \int \cdots \int dr_3 \cdots dr_{m+2} \sum_{\substack{m+2 \text{ molecules} \\ \text{connected to } r_1 \text{ and } r_2 \\ \text{even when } r_1 \text{ and } r_2 \text{ are erased}}} H(r_{13}, \cdots, r_{1,m+2}, r_{23}, \cdots, r_{2,m+2})$$

Sum over all products for which each molecule of the set $\{r_3, \cdots, r_{m+2}\}$ is a junction and is connected to r_1 and r_2 by an independent path and also the molecules $\{r_3, \cdots, r_{m+2}\}$ are connected with each other even when we erase the lines ending at r_1 and r_2 .

§ 5. Another possible expansion schemes

We have, in the preceding sections, obtained the formulae for the free energy and the radial distribution function in the hyper-netted chain approximation by grouping the graphs consisting of lines corresponding to $f^{(n)}$, according to the times of identification needed to reduce them to a line or a ring. When $\phi(r)$ has the Fourier transform, we can do the same thing by first, expanding $f^{(n)}(r)$ in (4) in powers of $-\phi(r)/kT$, then corresponding a line to a factor $-\phi(r)/kT$ and again grouping the graphs by the times of identification needed to reduce them to a line or a ring; which leads to another recurrence equations in the same approximation. Moreover, we can do the same thing by dividing the potential into two parts as

$$\phi(r) = \phi_1(r) + \phi_2(r), \quad (38)$$

$\phi_2(r)$ being assumed to have the Fourier transform, then introducing $f_1(r)$ and $j(r)$ by

$$e^{-\phi_1(r)/kT} - 1 = f_1(r), \quad -\frac{\phi_2(r)}{kT} = j(r) \quad (39)$$

and expanding as

$$e^{-\phi(r)/kT} - 1 = f_1(r) + \sum_{n=1}^{\infty} \frac{1}{n!} j(r)^n + f_1(r) \sum_{n=1}^{\infty} \frac{1}{n!} j(r)^n, \quad (40)$$

then corresponding a line to $f_1(r)$ and $j(r)$ and again grouping the graphs by the times of identification needed to reduce them to a line or a ring. With this expansion, what we are to evaluate are (1) with (2) or (4) and (33) with (34), with $f^{(0)}(r)$ replaced by the right-hand side of (40). The merit of this method is in that the first Fourier transformation may be done analytically; while the second, third, ... Fourier transformations will have to be performed in the numerical way.

The contributions to the free energy are calculated separately for each group of the graphs grouped as in Fig. 16, as follows:

$$\left(-\frac{A'}{NkT}\right)_b = -\frac{\rho}{2} F^{(0)}(0) \quad (41)$$

with

$$f^{(0)}(r) = f_1(r) + j(r). \quad (42)$$

$$\left(-\frac{A'}{NkT}\right)_{b+c} = \frac{1}{2\rho V} \sum_k \left\{ -\ln(1 - \rho F^{(0)}(k)) - \rho F^{(0)}(k) - \frac{1}{2} \rho^2 F^{(0)}(k)^2 \right\}. \quad (43)$$

$$\begin{aligned} \left(-\frac{A'}{NkT}\right)_{b+c+d} &= \frac{\rho}{2} \int d\mathbf{r} f_1(r) j(r) + \frac{\rho}{4} \int d\mathbf{r} j(r)^2 \\ &+ \frac{\rho}{2} [F^{(1)}(0) - F^{(0)}(0)] - \frac{\rho}{2} \int d\mathbf{r} f_1(r) [h^{(0)}(r) + j(r)] \\ &- \frac{\rho}{4} \int d\mathbf{r} [h^{(0)}(r) + j(r)]^2 \\ &= \frac{\rho}{2} [F^{(1)}(0) - F^{(0)}(0)] - \frac{\rho}{2} \int d\mathbf{r} f^{(0)}(r) h^{(0)}(r) - \frac{\rho}{4} \int d\mathbf{r} h^{(0)}(r)^2 \end{aligned} \quad (44)$$

with

$$H^{(0)}(k) = \frac{\rho F^{(0)}(k)^2}{1 - \rho F^{(0)}(k)} \quad (45)$$

and

$$f^{(1)}(r) = f_1(r) e^{h^{(0)}(r) + j(r)} + e^{h^{(0)}(r) + j(r)} - 1 - h^{(0)}(r). \quad (46)$$

$$\left(-\frac{A'}{NkT}\right)_{R_c} = \left(-\frac{A'}{NkT}\right)_{R_b} - \left(-\frac{A'}{NkT}\right)_{R_b} - \frac{1}{2} \int d\mathbf{r} h''(r) [f^{(1)}(r) - f^{(2)}(r)], \quad (47)$$

and so forth. As the result, if we make the correspondence $B^{(0)}=b$, $B^{(1)}=b'-c''$, $B^{(2)}=d''$,... and $R^{(0)}=R_b$, $R^{(1)}=R_c$, $R^{(2)}=R_d$, ..., we have eqs. (24)–(28) with the recurrence equations:

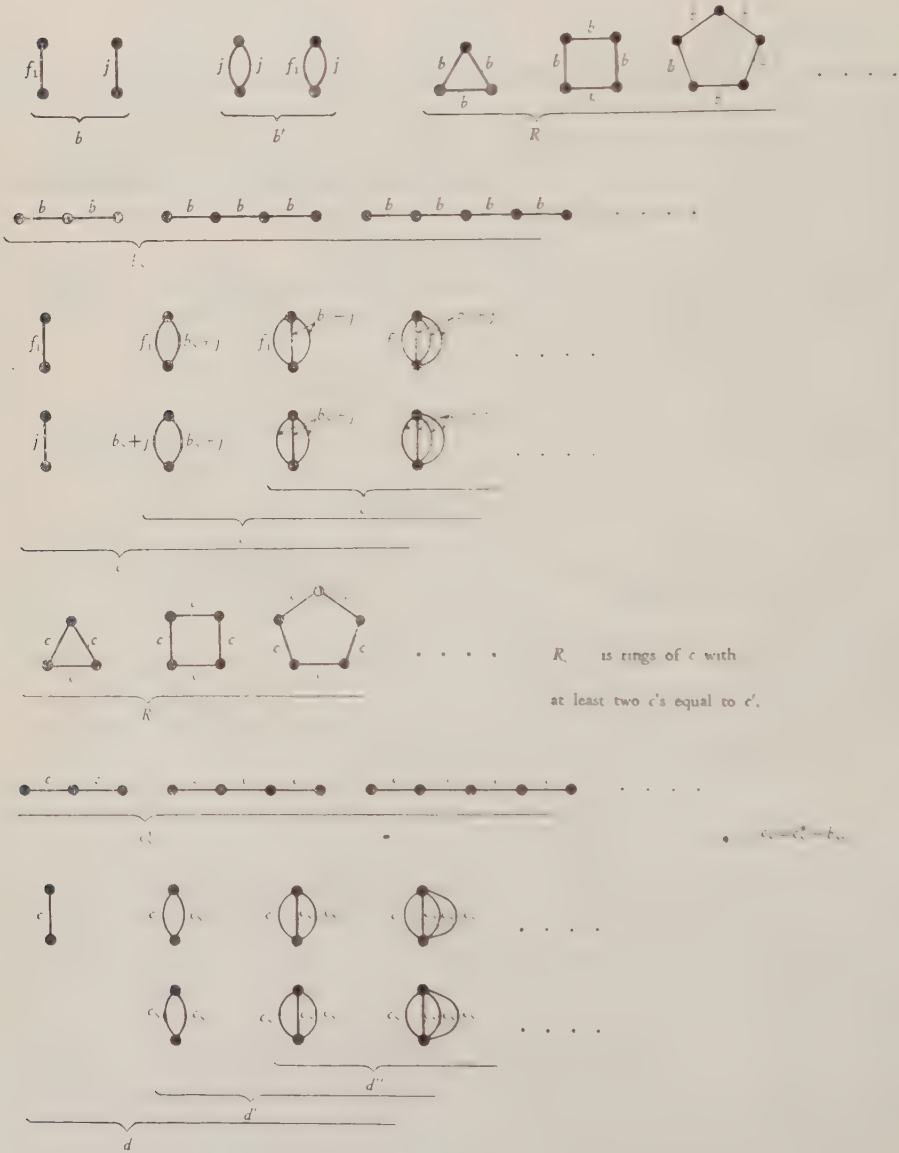


Fig. 16. b ; b' ; R_b ; b_s , c'' ; c , c' , R_c^* , R_c ; c_s^* , c_s , d'' ; d , d' ,...

$$\begin{aligned}
 H^{(n)}(k) &= \frac{\rho F^{(n)}(k)^2}{1 - \rho F^{(n)}(k)} \\
 f^{(n+1)}(r) &= f^{(n)}(r) e^{\Delta h^{(n)}(r)} + e^{\Delta h^{(n)}(r)} - 1 - \Delta h^{(n)}(r), \quad n \geq 1; \\
 f^{(1)}(r) &= f_1(r) e^{h^{(0)}(r) + j(r)} + e^{h^{(0)}(r) + j(r)} - 1 - h^{(0)}(r) \\
 f^{(0)}(r) &= f_1(r) + j(r);
 \end{aligned} \tag{48}$$

$F^{(n)}(k)$ and $H^{(n)}(k)$ being the Fourier transforms of $f^{(n)}(r)$ and $h^{(n)}(r)$, respectively. As before, we put

$$\mathcal{F}(r) = \lim_{n \rightarrow \infty} f^{(n)}(r), \quad \mathcal{H}(r) = \lim_{n \rightarrow \infty} h^{(n)}(r), \quad \mathcal{J}(k) = \lim_{n \rightarrow \infty} F^{(n)}(k). \tag{48'}$$

Then, the free energy in the hyper-netted chain approximation is given by eq. (29) again.

We get for the radial distribution function

$$g(r) = \exp[-\phi(r)/kT + \mathcal{H}(r)] \tag{49}$$

in the analogous way as in the previous section.*

To compare with the results obtained by various authors in the past for fluids of particles interacting by the Coulomb potential with and without hard core in a compensating field, we write the results in the ring and the watermelon approximation. In the ring approximation,**

$$\begin{aligned}
 \left(-\frac{A'}{NkT}\right)_{b+b'+R_b} &= \frac{\rho}{2} F^{(0)}(0) + \frac{\rho}{2} \int d\mathbf{r} f_1(r) j(r) + \frac{\rho}{4} \int d\mathbf{r} j(r)^2 \\
 &+ \frac{1}{2\rho V} \sum_k \left\{ -\ln(1 - \rho F^{(0)}(k)) - \rho F^{(0)}(k) - \frac{1}{2} \rho^2 F^{(0)}(k)^2 \right\}.
 \end{aligned} \tag{50}$$

In the watermelon approximation***

$$\begin{aligned}
 \left(-\frac{A'}{NkT}\right)_{b+b'+R_b+c''} &= \left(-\frac{A'}{NkT}\right)_{b+b'+R_b} \\
 &+ \frac{\rho}{2} \int d\mathbf{r} \left\{ f_1(r) [e^{h^{(0)}(r) + j(r)} - 1 - h^{(0)}(r) - j(r)] \right. \\
 &\left. + e^{h^{(0)}(r) + j(r)} - 1 - h^{(0)}(r) - j(r) - \frac{1}{2} (h^{(0)}(r) + j(r))^2 \right\}
 \end{aligned} \tag{51}$$

$$\begin{aligned}
 &= \frac{\rho}{2} F^{(1)}(0) + \frac{1}{2\rho V} \sum_k \left\{ -\ln(1 - \rho F^{(0)}(k)) - \rho F^{(0)}(k) - \frac{1}{2} \rho^2 F^{(0)}(k)^2 \right\} \\
 &- \frac{\rho}{2} \int d\mathbf{r} f^{(0)}(r) h^{(0)}(r) - \frac{\rho}{4} \int d\mathbf{r} h^{(0)}(r)^2.
 \end{aligned} \tag{51'}$$

* The graphs which contribute to $g(r)$ are represented by Fig. 13 if we replace the large letters B, C, D, \dots by the small letters b, c, d, \dots .

** Note that, when $f_1(r) = 0$, the third term of the right-hand side cancels with the sum of the last term in $\{ \}$.

*** Note that, when $f_1(r) = 0$, the $h^{(0)}(r) + j(r)$ is the cut-off Coulombic potential if $j(r)$ is Coulombic.

The ring approximation without hard core, eq. (50) with $f_i(r)=0$, is equal to Debye-Hückel's result as was shown by Mayer,⁹⁾ which was also obtained by Zubarev¹⁰⁾ by the use of collective variables. Mayer⁹⁾ took account of the terms up to the first power of $f_i(r)$ in eq. (51), and Haga¹¹⁾ considered some terms in addition. Yukhnovsky¹²⁾ has given the formulae in the watermelon approximation up to the first order of $f_i(r)$ with some additional terms representing a triangle of watermelons, by the use of the collective variables. It is to be noted that our formulae have been obtained for one-component system, while the theories compared⁹⁻¹²⁾ are valid also for multi-component systems.

§ 6. Conclusion

We have thus found the possibility to compute the free energy and the radial distribution function of fluids, in the hyper-netted chain approximation, with account of far more graphs than those considered in the approximations in the past.

The numerical computations for fluids of particles with Lennard-Jones potential and those of hard spheres and the extension of our method of approximation to the multi-component systems will be given in the following papers.

Acknowledgement

I wish to express my cordial thanks to Mr. R. Abe. He showed me the result of his investigation before publication, which has stimulated me to make investigations into these problems.

References

- 1) J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (Wiley, 1940) Chapter 13.
- 2) E. W. Montroll and J. E. Mayer, *J. Chem. Phys.* **9** (1941), 626.
- 3) H. S. Green, *Proc. Roy. Soc. A***189** (1947), 103.
- 4) K. Hiroike, to be published in *J. Phys. Soc. Japan* **13** (1958), No. 12.
G. S. Rushbrooke and H. I. Scoins, *Proc. Roy. Soc. A***216** (1953), 203.
- 5) R. Abe, to be published in *J. Phys. Soc. Japan*.
- 6) J. E. Mayer and E. W. Montroll, *J. Chem. Phys.* **9** (1941), 2.
- 7) N. N. Bogolyubov, *Problems of the Dynamical Theory in Statistical Physics* (1946).
- 8) K. Hiroike, *J. Phys. Soc. Japan* **12** (1957), 864.
- 9) J. E. Mayer, *J. Chem. Phys.* **18** (1950), 1426.
E. Meeron, *J. Chem. Phys.* **26** (1957), 804.
- 10) D. N. Zubarev, *Dokl. Akad. Nauk SSSR* **95** (1954), 757.
- 11) E. Haga, *J. Phys. Soc. Japan* **8** (1953), 714.
- 12) I. R. Yukhnovsky, *J. Exp. Theor. Phys.* **34** (1958), 379.
- 13) E. Meeron, *J. Chem. Phys.* **28** (1958), 630.
- 14) E. Meeron, *Phys. Fluids* **1** (1958), 139.

Meson Production in Meson-Nucleon Interaction

S. K. KUNDU

Department of Theoretical Physics

Indian Association for the Cultivation of Science, Calcutta-32, India

(Received June 7, 1958)

The meson production by mesons in the low-energy limit has been re-examined in the frame of static model of the extended source theory. Using Feynman's technique the calculation has been performed up to fifth order in the coupling constant. The renormalization procedure is the same as developed and applied by Chew in meson-nucleon scattering processes. To reduce the Feynman diagram an approximation similar to that of the Tamm-Dancoff type has been used. The total cross section for charged and neutral pions from the pion-nucleon collision in $T=3/2$, $J=3/2$ state are obtained for incident meson energy 260 to 570 Mev. The theoretical results have been compared with those of Franklin, Barshay and others. Franklin's result that production of charged pions is higher than that of neutral pions in the cut-off theory is confirmed. The present calculation predicts the charged pion production of about 7 mb at 500 Mev against the available experimental cross section of a maximum of 10 mb at the same energy.

§ 1. Introduction

In view of the recent experimental evidences¹⁾ of the production of one additional pion in pion-nucleon scattering at low energy, the problem of double scattering of π -mesons by nucleons has of late been approached by Barshay²⁾ and Franklin³⁾. They have made use of the extended fixed source meson theory of Chew⁴⁾ et al, and based their calculations on a new method for meson-nucleon scattering developed by Wick⁵⁾ which involves the study of the eigenstates of the meson-nucleon Hamiltonian. Franklin has calculated the transition matrix from one meson state to two meson state using two meson eigenstate of the total meson-nucleon Hamiltonian analogous to ordinary meson-nucleon scattering, which needs the one meson eigenstate to describe the transition matrix of the scattering process. Barshay also, using a similar procedure (as studied by Low⁶⁾), has obtained results which are quite different from those obtained by Franklin. The discrepancy between the results of Franklin and Barshay can be seen by comparing Franklin's equation (27) with Barshay's equation (20) for the production T -matrix. It is noted that apart from a factor $1/1/2$ (due to different procedure of normalization) Barshay has obtained only twice the first term of (27), the second term of (27), however, vanishes for the $3/2$ isotopic spin state and the rest of the discrepancy as has been pointed out by Franklin is due to the fact that Barshay has used different approximation to identify the certain parts of the production T -matrix with the elastic scattering T -matrix.

In this work we have adopted the Born-approximation method and accordingly started from the lowest order Feynman's diagram, for the production of mesons in meson-nucleon interaction. To obtain higher order corrections to the lowest order diagrams, Dyson's procedure of modification of nucleon lines and vertex operators are used. However, one must be a little careful in estimating the order of magnitudes of the successive terms of the Born-approximation. The validity of this perturbation expansion was checked by Friedman⁷ and Kundu⁸ in a different problem, namely the anomalous magnetic moment of the nucleon. Again in the present problem it is found that the 5th order correction to the production cross section is small compared with the 3rd order one and probably the higher order corrections are quite negligible. It must be mentioned that though the unitarity of the S -matrix is lost in such an expansion but within the limit of the static-source approximation the theory does not predict the cross-section which is beyond the geometrical one. As is apparent that the number of Feynman diagrams will increase enormously while calculating the 5th order correction to the lowest order Feynman diagrams of the production process, we simplify our calculations by restricting ourselves to a Tamm-Dancoff type of approximation⁹. If, in any diagram, a line drawn perpendicular to the trajectory of the nucleon cuts more than three mesons (real and virtual) at an instant of time, we neglect the diagram.

In the present calculation we have determined the production of an additional pion (π^+ and π^0) in the pion-nucleon scattering, at the energy range 260 Mev to 570 Mev. The theoretically calculated values of the pion production cross section in millibarns at different energies is given in tabular form in the last section, and the same was compared over with those of Franklin and Barshay. Rodberg¹⁰ and Kazes¹¹, using a 'one-meson approximation', have obtained an integral equation of Low.¹² Some discrepancies in the results of Barshay and Franklin have been cleared. Their results are in conformity with our calculations based on pure perturbation analysis.

It is worth-noting¹³, in this connection, that this problem of double scattering was also treated in the relativistic formalism with the use of both the covariant way of Feynman as well as in the conventional perturbation theory. But this calculations were restricted in the ps-ps theory only; and, as is well known, the production cross-section was much too small to predict the experimental results, even with a high value of the coupling constant.

§ 2. Matrix element for the production process

Following Chew, we write the lowest order matrix element M for the meson production corresponding to Fig. I, as follows:

$$M(p^*, q^*, k) = M_a(p^*, q^*, k) + M_b(p^*, q^*, k) + M_c(p^*, q^*, k) \\ + \text{similar terms interchanging } p \text{ and } q. \dots \dots (1)$$

¹³ The author is grateful to Professor G. F. Chew for informing him the results of these important papers and also for valuable suggestions which led to the material development of this paper.

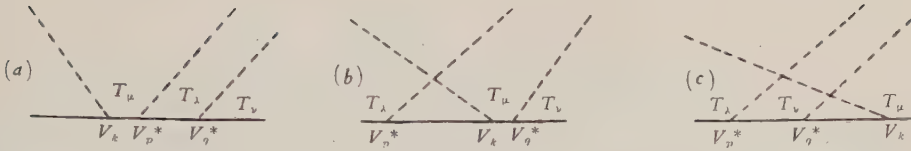


Fig. 1. Feynman diagrams in the first Born-approximation: Other three diagrams are obtained by interchanging p and q . (All T should be replaced by τ .)

with

$$M_a = V_q^* S(E - \omega_q) V_p^* S(E - \omega_q - \omega_p) V_k$$

$$M_b = V_q^* S(E - \omega_q) V_k S(E - \omega_q + \omega_k) V_p^*$$

and

$$M_c = V_k S(E + \omega_k) V_q^* S(E + \omega_k - \omega_q) V_p^*.$$

In writing down the above matrix elements we have used the interaction Hamiltonian $H_I = \sum_k (a_k V_k + a_k^* V_k^*)$, where a_k 's are the usual creation and annihilation operators $V_k = i f (4\pi)^{1/2} t^{-1} (2\omega_k)^{-1/2} (\sigma \cdot k) \tau_u v(k)$, and $S(E)$ denotes as usual the nucleon propagator in the sense of Chew. Here-in-after the argument in the parenthesis will be dropped for brevity.

To obtain the fifth order correction term we apply the usual renormalization procedure of Dyson as formulated by Chew for the static-source theory. We denote this contribution by $M^{(r)}$, etc., where

$$M^{(r)} = M_a^{(r)} + M_b^{(r)} + M_c^{(r)} + \text{terms interchanging } p \text{ and } q \dots \dots \dots (2)$$

The $M_a^{(r)}$'s are given by

$$M_a^{(r)} = V_{qr}^* S_r(E - \omega_q) V_{pr}^* S_r(E - \omega_q - \omega_p) V_k$$

where

$$V_{qr}^* = V_q^* \left[1 + 1/9 \sum_l V_l V_l^* \left(\frac{1}{E_2 - \omega_l} \frac{1}{E_1 - \omega_l} - \frac{1}{\omega_l^2} \right) \right]$$

and

$$S_r^{-1}(E) = S^{-1}(E) - \sum_l V_l V_l^* \left(\frac{1}{E - \omega_l} + \frac{1}{\omega_l} + \frac{E}{\omega_l^2} \right)$$

are renormalized vertex and nucleon propagators with ω_l being the virtual meson energy and E_1, E_2 are the nucleon energies before and after the vertex.

The net contribution up to the fifth order in the coupling constant from above is given by

$$M + M^{(r)}$$

where the prime in the second term denotes that only the 5th order contributions from $M^{(r)}$ has been taken.

In addition, we are left with irreducible diagrams, a sample diagram of which has shown

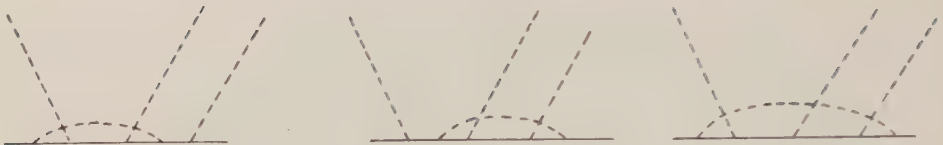


Fig. II. $M_a^{(i)}$

in Fig. II.

$$M^{(5)} = M_a^{(5)} + M_b^{(5)} + M_c^{(5)} + \text{terms interchanging } p \text{ and } q, \dots\dots\dots (3)$$

where $M_a^{(5)}$'s are the corresponding contributions from the 5th order irreducible diagrams obtained from the diagram of the type 'a'. It should be pointed out that the part $M^{(5)}$ does not contribute to $M^{(5)}$ because of our assumption that diagrams in which there are more than three mesons present at an instant of time are to be neglected. We have for $M_a^{(5)}$

$$M_a^{(5)} = \sum_i (f_1 V_q^* V_l V_p^* V_k V_l^* + f_2 V_l V_q^* V_p^* V_l^* V_k + f_3 V_l V_q^* V_p^* V_k V_l^*) \\ + \text{similar terms interchanging } p \text{ and } q.$$

and $M_b^{(5)}$ is obtained by using Goldberger's crossing symmetry property. f_1, f_2 etc., are functions of ω arising from the energy denominator of the nucleon propagator.

Thus the total 5th order contribution to the matrix element $T(p^*, q^*, k)$ for the production of two mesons of momenta p and q from an incident meson of momentum k is thus given by

$$T(p^*, q^*, k) = M + M^{(5)} + M^{(6)}.$$

Collecting all the contributions from (1), (2) and (3), we can write $T(p^*, q^*, k)$ in the following simple form

$$T(p^*, q^*, k) = \frac{f^3}{\mu^3} (4\pi)^{3/2} \frac{N^*{}_{\lambda}{}^*(p) N^*{}_{\mu}{}^*(q) N_{\nu}(k)}{\sqrt{2\omega_p} \sqrt{2\omega_q} \sqrt{2\omega_k}} \left[\sum^{p,q} F_{\lambda\mu\nu, \sigma\tau\eta}(\omega) \tau_{\lambda\mu\nu, \sigma\tau\eta} \right] \quad (4)$$

where $\sum^{p,q}$ denotes the summation of all terms obtained from the all possible permutations of the indices of τ and σ ; $\tau_{\lambda\mu\nu}$ stands for $\tau_\lambda \tau_\mu \tau_\nu$ and $\sigma_{\sigma\tau\eta}$ for $(\sigma \cdot p) (\sigma \cdot k) (\sigma \cdot q)$. Again $F_{\lambda\mu\nu, \sigma\tau\eta}(\omega)$ are functions of the energy variables $\omega_p, \omega, \omega_q$ and also of the meson mass μ . Of the total number of terms in the expression (4) half of the terms are obtained by interchanging p and q . The detailed expressions for $F_{\lambda\mu\nu, \sigma\tau\eta}(\omega)$ will be given later for the particular case of production of two mesons in isotopic spin-state $T=2$ and $T=1$ from the initial meson-nucleon $T=3/2$ and $J=3/2$.

It may be pointed out that due to our method of approximation we have vanishing contribution from the energy co-efficients $F_{\mu\lambda\nu, \sigma\tau\eta}(\omega)$, $F_{\lambda\mu\nu, \sigma\tau\eta}(\omega)$, $F_{\mu\lambda\nu, \sigma\tau\eta}(\omega)$ and $F_{\mu\lambda\nu, \sigma\tau\eta}(\omega)$ appearing in (4) together with another four terms obtained respectively by the interchange of indices $\lambda\mu$ of the first two and $\mu\nu$ of the last two co-efficients.

In the following we shall be interested only in the specific determination of the π^+ and π^0 production from the meson-nucleon interaction in the isotopic spin-state, T_{33} . We note here that due to the conservation of isotopic spin the matrix element M should

satisfy

$$\langle T', T'_3 | M(T, T_3) | T, T_3 \rangle = M(T, T_3) \delta_{TT'} \delta_{T_3 T'_3}$$

where the initial and final states are characterised by the quantum numbers T , T_3 and T' , T'_3 .

First, we calculate the additional production of π^+ -meson in the reaction $P + \pi^+ \rightarrow N + \pi^+ + \pi^+$ and we abbreviate the corresponding matrix element by I

$$\langle I_{22} | T(p^*, q^*, k) | T_{33} \rangle \quad (5)$$

where T denotes the total meson-nucleon iso-spin states and I denotes the total iso-spin states of the two outgoing mesons. T_3 and I_3 are the third component of T and I respectively. Similarly the matrix element

$$\langle I_{21,11} | T(p^*, q^*, k) | T_{33} \rangle \quad (6)$$

denotes the production of π^0 in $P + \pi^+$ scattering.

For the process given in (5) we have from (4) the following expressions:

$$\sum^P \sigma_{pqk} \Gamma_{pqk}(\omega), \quad (7)$$

and for the process given in (6)

$$\sum^P \sigma_{pqk} \Gamma'_{pqk}(\omega) \quad (8)$$

where

$$\begin{pmatrix} \Gamma_{pqk} & \Gamma'_{pqk} \\ \Gamma_{pkq} & \Gamma'_{pkq} \\ \Gamma_{qkp} & \Gamma'_{qkp} \\ \Gamma_{kpq} & \Gamma'_{kpq} \\ \Gamma_{kqp} & \Gamma'_{kqp} \\ \Gamma_{qp k} & \Gamma'_{qp k} \end{pmatrix} = \sqrt{2} \begin{pmatrix} (F_{\lambda\mu\nu,pqk} + F_{\nu\mu\lambda,pqk}) & \sqrt{2} (F_{\lambda\mu\nu,pqk} - F_{\lambda\nu\mu,pqk} + F_{\nu\lambda\mu,pqk}) \\ (F_{\lambda\mu\nu,pkq}) & \sqrt{2} (F_{\lambda\mu\nu,pkq} - F_{\lambda\nu\mu,pkq} + F_{\nu\lambda\mu,pkq}) \\ (F_{\nu\mu\lambda,qkp}) & \sqrt{2} (-F_{\lambda\nu\mu,qkp} + F_{\nu\lambda\mu,qkp}) \\ (F_{\lambda\mu\nu,kpq}) & \sqrt{2} (F_{\lambda\mu\nu,kpq} - F_{\lambda\nu\mu,kpq} + F_{\nu\lambda\mu,kpq}) \\ (F_{\nu\mu\lambda,kqp}) & \sqrt{2} (-F_{\lambda\nu\mu,kqp} + F_{\nu\lambda\mu,kqp}) \\ (F_{\lambda\mu\nu,qpk} + F_{\nu\mu\lambda,qpk}) & \sqrt{2} (F_{\lambda\mu\nu,qpk} - F_{\lambda\nu\mu,qpk} + F_{\nu\lambda\mu,qpk}) \end{pmatrix} \quad (9)$$

The values of the matrix element appearing in the right-hand side of equation (9) are given in the Appendix. We note here that in the production of $2\pi^+$ and $\pi^+\pi^0$ from $P\pi^+$ system the former is symmetric with respect to the interchange of the isotopic variables of the two outgoing mesons while the latter one is antisymmetric.

In deducing (7) and (8) we have replaced the τ -matrices by their eigenvalues for the respective transitions from the given initial isotopic spin states to the final isotopic spin states as denoted in the expressions (5) and (6). This replacement of the τ -matrices by their proper-value can be effected by noting that one can use τ_+ , τ_- and τ_3 instead of τ_1 , τ_2 , τ_3 (τ_μ , etc). This requires only a slight modification that since τ_+ , τ_- are not hermitian one should be careful about the 'star' in the emission. One should remember also that

$$\tau_- P = \frac{1}{\sqrt{2}} N, \quad \tau_+ N = \frac{1}{\sqrt{2}} P, \quad \tau_- N = \tau_+ P = 0,$$

$$\tau_3 P = P, \quad \tau_3 N = -N,$$

where N and P denote the neutron and proton states of the nucleon respectively.

Following this, one can at case construct a table of the eigenvalues of different combinations of the three τ -matrices (i.e. of τ_z , τ_x , τ_y , etc.) from different transitions such as $P + \pi^+ \rightarrow N + \pi^+ + \pi^+$, etc., just similarly as obtained by Wick (cf. Table 1, page 348 ref. no. 5) for the combination of two τ -matrices in transitions $P + \pi^+ \rightarrow P + \pi^+$, etc.

We now expand the scattering matrix $T(p^*, q^*, k)$ in terms of the projection operators for the various states of the total angular momentum. We keep here only the angular momentum states characterised by $J=3/2$. Using the well-known projection operators P_3 for the state $J=3/2$, we can write the T -matrix for the production of π^- from equation (6) in the following way,

$$\begin{aligned} & \langle J'_{3/2}; I_{21,11} | T(p^*, q^*, k) | T_{3/2}; J_{3/2} \rangle \\ &= (1/3) [(\sigma \cdot p) \{ \Gamma_{pkq} [2(k \cdot q) + i\sigma \cdot (k \times q)] \} + \{ \Gamma_{kqp} [2(k \cdot q) + i\sigma \cdot (k \times q)] (\sigma \cdot P) \} \\ &+ \text{terms interchanging } p \text{ and } q. \end{aligned} \quad (10)$$

The result for the square of the matrix element averaged and summed over initial and final spin states is given by

$$|T_{\pi^-}(p^*, q^*, k)|^2 = \int_{\mu^2}^{\mu_0^2} (4\pi)^2 \frac{1}{8\omega_p \omega_q \omega_k} \left\{ \left(\frac{4\pi}{27} \right) p^2 q^2 k^2 (a_1 + b_1 \cos^2 \theta) \right\} \quad (11)$$

where $T_{\pi^-}(p^*, q^*, k)$ stands for the T -matrix element appearing in the left-hand of the above equation (10).

In a similar way the square of the matrix element, $T_{\pi^0}(p^*, q^*, k)$ for the production of π^0 is given by

$$|T_{\pi^0}(p^*, q^*, k)|^2 = \int_{\mu^2}^{\mu_0^2} (4\pi)^2 \frac{1}{8\omega_p \omega_q \omega_k} \left\{ \left(\frac{4\pi}{27} \right) p^2 q^2 k^2 (a_2 + b_2 \cos^2 \theta) \right\} \quad (12)$$

where

$$\begin{aligned} a_1 = & \{ 2(\Gamma_{pkq} + \Gamma_{kqp}) + (\Gamma_{qkp} - \Gamma_{kpq}) \}^2 + \{ (\Gamma_{kqp} - \Gamma_{pkq}) \\ & + (\Gamma_{kpq} - \Gamma_{qkp}) \}^2 + \{ (\Gamma_{qkp} + \Gamma_{kpq}) - (\Gamma_{pkq} + \Gamma_{kqp}) \}^2 \end{aligned}$$

and

$$\begin{aligned} b_1 = & 3 \{ 2(\Gamma_{qkp} + \Gamma_{kpq}) + (\Gamma_{pkq} - \Gamma_{kqp}) \}^2 - \{ (\Gamma_{qkp} + \Gamma_{kpq}) - (\Gamma_{pkq} + \Gamma_{kqp}) \}^2 \\ & + 2[\{ 2(\Gamma_{pkq} + \Gamma_{kqp}) + (\Gamma_{qkp} - \Gamma_{kpq}) \} \{ 2(\Gamma_{qkp} + \Gamma_{kpq}) + (\Gamma_{pkq} - \Gamma_{kqp}) \} \\ & + \{ 2(\Gamma_{qkp} + \Gamma_{kpq}) + (\Gamma_{pkq} - \Gamma_{kqp}) \} \{ (\Gamma_{kqp} - \Gamma_{pkq}) + (\Gamma_{kpq} - \Gamma_{qkp}) \} \\ & + \{ (\Gamma_{kqp} - \Gamma_{pkq}) + (\Gamma_{kpq} - \Gamma_{qkp}) \} \{ 2(\Gamma_{pkq} + \Gamma_{kqp}) + (\Gamma_{qkp} - \Gamma_{kpq}) \}]. \end{aligned} \quad (13)$$

a_2 and b_2 are obtained from the same expressions for a_1 and b_1 respectively given above where now all the Γ 's should be replaced by their primes.

§ 3. Cross section

The differential cross section for the production of an additional pion in the scattering of a pion of momentum k into two types of pions of momenta p and q is given by

$$d\sigma = 2\pi/v \sum \rho_F |T(p^*, q^*, k)|^2 \delta(E_i - E_f) \quad (14)$$

where v = relative velocity between the incident meson and nucleon and in the static limit $v = k/\omega_k$ and ρ_F is the number of states per unit energy level. Remembering the conservation of energy, we get in our approximation $\omega_k = \omega_p + \omega_q$. In (14) we will use the value of $T_{\pi^+}(p^*, q^*, k)$ and $T_{\pi^0}(p^*, q^*, k)$ for $T(p^*, q^*, k)$ for the production cross section of π^+ and π^0 respectively. Using the values of (11) and (12) together with (13), we get from (14) that the total cross section σ for the production of π^+ and π^0 in $P + \pi^+ \rightarrow N + \pi^+ + \pi^+$ and $P + \pi^+ \rightarrow P + \pi^+ + \pi^0$, respectively, is given by

$$\sigma = \frac{1}{3\pi} \frac{f^6}{\mu^6} (\omega_k^2 - \mu^2)^{1/2} \int d\Omega \int_{\mu}^{\omega_k - \mu} (\omega_p^2 - \mu^2)^{3/2} [(\omega_k - \omega_p)^2 - \mu^2]^{3/2} (a_i + b_i \cos^2 \theta) d\omega_p \quad (15)$$

where $i=1$ denotes the production of π^+ and $i=2$ for π^0 production.

§ 4. Results and discussions

We now give the numerical results for the total cross section σ , determined from the formula (15). The results are tabulated for incident meson energies (laboratory system) from 260–570 Mev and compared with those of Franklin and Barshay. In obtaining the results two adjustable parameters, the coupling constant f^2 and the cut-off

Table 1.

Processes	Incident meson energy in Mev.	Cross section in mb.		
		Present Calculation	Results of Franklin	Results of Barshay
$P + \pi^+ \rightarrow N + \pi^+ + \pi^+$	260	0.067	0.040	≈ 0.071
	360	0.889	0.970	≈ 0.860
	570	15.640	9.250	
	260	0.037	0.037	≈ 0.440
$P + \pi^+ \rightarrow P + \pi^+ + \pi^0$	360	0.484	1.040	≈ 0.550
	570	8.017	6.540	

energy ω_c are chosen to be .093 and 6μ respectively. It appears from the table below that the charged pion production is larger than the neutral pion production, contrary to what has been obtained by Barshay. The only experimental data of Blau and Caulton⁽¹³⁾ on the charged pion production cross section by free nucleon at incident 500 Mev meson energy give a rough estimate of 3.5 to 10 mb. In the present calculation we find on using the above values of the coupling constant and the cut-off energy the production cross section of 7 mb at 500 Mev meson energy. This is fairly in good agreement with their estimate. Calculations at still higher energy such as 570 Mev predicts a value of 15 mb, of the production cross section indicating a complete breakdown of the cut-off theory and in fact this result exceeds the geometrical cross section ($\pi\lambda^2$). However, it

must be pointed out that the energy for which the cross section goes beyond the geometrical one is too high to expect any fruitful result with the static-source theory. Another point must be borne in mind that assumed perturbation expansion destroys the unitarity of the S -matrix leading to such catastrophe. But in support of the validity of our calculation we note that the 5th order terms are small compared with the 3rd order ones and probably the 7th order terms are still smaller than the 5th. It may be remarked that the present procedure is not exact. To discuss fully one should use the dispersion relation for pion-nucleon scattering in a similar way as introduced by Miyazawa¹⁴ in the investigation of nuclear forces and magnetic moment problems.

Acknowledgement

I wish to express my sincere appreciation to Dr. S. N. Biswas, for suggesting this problem and for his continued assistance. Thanks are also due to Dr. N. C. Sil for helpful discussion

Appendix

$$\begin{aligned}
 F_{\lambda\mu\nu, kpg} &= 2\omega_f \omega_c^{-1} \omega_q^{-1} \\
 F_{\lambda\mu\nu, pkeq} &= -\omega_p^{-1} \omega_q^{-1} - \omega_f [\omega_p^{-1} \{(\omega_c + \omega_q)^{-1} - 7\omega_c^{-1} + 4(\omega_c - \omega_k)^{-1}\} + \omega_q^{-1} \\
 &\quad \times \{7(\omega_c + \omega_p)^{-1} - \omega_c^{-1}\}] \\
 F_{\lambda\mu\nu, pqk} &= -2\omega_f \omega_p^{-1} [(\omega_c - \omega_k)^{-1} + (\omega_c + \omega_q)^{-1}] \\
 F_{\lambda\mu\nu, qpk} &= -2\omega_f \omega_c^{-1} (\omega_c + \omega_q)^{-1} \\
 F_{\lambda\nu\mu, kpg} &= -2\omega_f \omega_c^{-1} (\omega_c - \omega_k)^{-1} \\
 F_{\lambda\nu\mu, kqp} &= -4\omega_f \omega_c^{-1} [(\omega_c + \omega_p)^{-1} + (\omega_c - \omega_k)^{-1}] \\
 F_{\lambda\nu\mu, pkeq} &= -2\omega_f \omega_p^{-1} [(\omega_c - \omega_k)^{-1} + (\omega_c + \omega_q)^{-1}] \\
 F_{\lambda\nu\mu, qkp} &= -2\omega_f \omega_c^{-1} (\omega_c + \omega_p)^{-1} \\
 F_{\lambda\nu\mu, pqk} &= \omega_p^{-1} \omega_k^{-1} + \omega_f [\omega_p^{-1} \{7(\omega_c + \omega_k)^{-1} - (\omega_c - \omega_k)^{-1} - 4(\omega_c + \omega_q)^{-1}\} \\
 &\quad + \omega_k^{-1} \{7(\omega_c + \omega_p)^{-1} - 5\omega_c^{-1}\} + 4\omega_c^{-1} (\omega_c + \omega_p)^{-1}] \\
 F_{\lambda\nu\mu, qpke} &= -4\omega_f \omega_c^{-1} \omega_k^{-1} \\
 \text{where } \omega_f &= \frac{\omega_c(\omega_c - \mu)}{6\pi} f^2 / \mu^2.
 \end{aligned}
 \tag{A}$$

In above we have written in detail only half of the $F_{\lambda\cdots}$'s appearing in the right-hand matrix of equation (9); the rest of the functions can be written very easily by interchanging the suffixes p and q and as well as λ and ν . It is interesting to note that one needs only to interchange the suffixes p and q in the right-hand side of the expression for $F_{\lambda\cdots}$ in equation (A).

References

- 1) V. G. Zinov and S. M. Korentchenko, "Pion production by $\pi^- - p$ collisions near threshold" Joint Institute of Nuclear Research, U. S. S. R. See also ref. no. 13.
- 2) S. Barshay, Phys. Rev. **103** (1956), 1102.
- 3) J. Franklin, Phys. Rev. **105** (1957), 1101.
- 4) G. F. Chew, Phys. Rev. **94** (1954), 1748, 1755.
- 5) G. C. Wick, Rev. Mod. Phys. **27** (1955), 339.
- 6) G. F. Chew and F. E. Low, Phys. Rev. **101** (1956), 1570.
- 7) M. H. Friedman, Phys. Rev. **97** (1955), 1123.
- 8) S. K. Kundu, Proc. Phys. Soc. London **72** (1958), 49.
- 9) H. S. Green, Phys. Rev. **95** (1954), 548.
- 10) L. Rodberg, Phys. Rev. **106** (1957), 1090.
- 11) E. Kazes, Phys. Rev. **107** (1957), 1131.
- 12) M. Cini and L. A. Radicati, Nuovo Cimento **8** (1951) 542; Y. Miyachi. Prog. Theor. Phys. **12** (1954), 243.
- 13) M. Blau and M. Caulton, Phys. Rev. **96** (1954), 150.
- 14) H. Miyazawa, Phys. Rev. **104** (1956), 1740.

On Quantum Theory of Transport Phenomena

—Steady Diffusion—

Sadao NAKAJIMA

Physical Institute, Nagoya University, Nagoya

(Received August 17, 1958)

A general formulation is given to the quantum theory of steady diffusion. In seeking for a steady solution of Liouville's equation, the boundary condition is taken into account by requiring that the solution should lead to a given distribution of average density. The distribution is to be determined by macroscopic law of diffusion and macroscopic boundary condition.

The basic equation thus obtained has a form similar to Bloch's kinetic equation and reduces to the latter in the limit of a system of weakly interacting particles. This is shown by generalizing a damping theoretical expansion of Kohn and Luttinger.

It is found that the Einstein relation is valid only for the symmetric part of diffusion- and electric conductivity tensors, in agreement with Kasuya's suggestion.

§ 1. Introduction

Recently Kubo and others^{1),2)} have succeeded in formulating quantum statistical expressions for transport coefficients such as electric and thermal conductivities. These formulae are just as general and rigorous as, say, the familiar expression for the partition function $Z = \text{Tr}(\exp\{-H/kT\})$. In practice appropriate approximation should, of course, be made in evaluating transport coefficients. The point is, however, that the conventional Bloch equation is nothing else but the lowest order approximation in a damping theoretical treatment of dynamical motion (see § 6) and by no means the most general way of approaching the problem.

Now, in the case of mechanical disturbances such as an external electric field, derivation of these formulae has been rather simple. The mechanical disturbance is expressed as a definite perturbing Hamiltonian and the deviation from equilibrium caused by it can be obtained by perturbation theory. On the other hand, thermal disturbances such as density and temperature gradients cannot be expressed as a perturbing Hamiltonian in an unambiguous way. Accordingly in the previous paper,³⁾ use had to be made of Onsager's assumption that the average regression of spontaneous fluctuation follows the macroscopic laws. As a result, certain ambiguity has been left over, concerning galvano-magnetic effects caused by thermal disturbances.

In the present paper, a general formulation free from such an assumption will be given to the theory of thermal disturbances and, in order to show the basic idea, the case of steady diffusion will be discussed in detail. In this case, the usual Bloch equation takes the form in which the drift term due to the density gradient is balanced by the

collision term. Our purpose is then to generalize the equation so as to include systems of strongly interacting particles (electrons, phonons, impurities, etc.). In other words, we shall seek for a steady solution of the quantum theoretical Liouville equation

$$i[H, \rho] = 0, \quad (1.1)$$

where H is the Hamiltonian of our system and ρ the density matrix, taking $\hbar = 1$. The boundary condition should be such that a steady gradient of average density is established by attaching suitable sources at the boundary. But for such a boundary condition, (1.1) would lead to an equilibrium density matrix, microcanonical, canonical or grand canonical. The microscopic detail of the interaction with sources at the boundary, however, should not be essential to the law of diffusion as an intrinsic property of a large system. We need only to suppose that the solution of (1.1) should lead to a given distribution of average density. The distribution is in turn to be determined by the macroscopic law of diffusion together with a given *macroscopic* boundary condition. The law of diffusion itself is a consequence from our solution of (1.1), so that the method is self-consistent.

In this sense, the thermal disturbance is a constraint upon dynamical motion and in fact appears as an effective potential in our basic equation (§ 4). The equation takes a form similar to the Bloch equation and reduces to the latter in the limit of a system of weakly interacting particles. This will be shown by applying a damping theoretical expansion due to Kohn and Luttinger³⁾ (§ 6).

An important consequence of the theory is that the well-known Einstein relation is valid only for the symmetric part of diffusion- and electric conductivity tensors. In other words, as regards the Hall effect, the gradient of chemical potential is *not* equivalent to the electric field. The difference becomes appreciable at low temperatures and under strong magnetic fields. Such a difference has first been suggested by Kasuya and is confirmed far beyond doubt by the present theory (§ 5).

From technical points of view, the present theory is similar to the so-called method of pseudo-potential.⁴⁾ It may also be regarded as generalization of Enskog's classical kinetic theory.⁵⁾ Indeed, Matsubara⁶⁾ has once tried the theory of hydrodynamic properties from the latter point of view. It seems, however, that neither physical nor mathematical details has ever been examined by these previous authors.

§ 2. Local equilibrium

For definiteness, let us take a system of similar particles (say, electrons in a metal), in which there exists a steady gradient of average density. The number density of particles is represented by the operator

$$n(\mathbf{x}) = \sum_j \delta(\mathbf{x} - \mathbf{x}_j) = \Omega^{-1} \sum_k n_k e^{ik \cdot \mathbf{x}}, \quad (2.1)$$

where Ω is the volume of the system and $n_k = \sum_j \exp\{-ik \cdot \mathbf{x}_j\}$ should not be confused with occupation numbers. The average density is then given by

$$\text{Tr}(\rho n(\mathbf{x})) = \langle n(\mathbf{x}) \rangle. \quad (2.2)$$

As mentioned in § 1, we suppose that the right-hand side of (2.2) is given as a solution of the macroscopic diffusion equation, hence (2.2) is the condition for the solution of (1.1) to satisfy.

Let us describe the system in equilibrium by the grand canonical distribution

$$\rho_e = \exp\{\hat{\xi}N - \beta H\} / \text{Tr}(\exp\{\hat{\xi}N - \beta H\}), \quad (2.3)$$

where $\hat{\xi}$ is the chemical potential, β the temperature and N the operator representing the total number of particles and commutable with the Hamiltonian H . We shall assume that the system in equilibrium is homogeneous so that

$$\langle n(\mathbf{x}) \rangle_e \equiv \text{Tr}(\rho_e n(\mathbf{x})) = n \quad (2.4)$$

is constant. Hence (2.3), though stationary, does not satisfy (2.2) when the density gradient does exist.

A possibility of satisfying (2.2) is given by the so-called local equilibrium distribution

$$\rho_i = \exp\{\psi + \hat{\xi}N - \beta H + \sum_{k \neq 0} \hat{\xi}_k n_{-k}\}. \quad (2.5)$$

Here ψ is a normalization constant to make

$$\text{Tr}(\rho_i) = 1 \quad (2.6)$$

and the $\hat{\xi}_k (k \neq 0)$ represent the fluctuation of chemical potential in space. From (2.6) it can be easily seen that

$$\langle n_{-k} \rangle = -\partial\psi / \partial\hat{\xi}_k, \quad \hat{\xi}_{-k} = -\partial S / \partial\langle n_k \rangle \quad (2.7)$$

where the entropy S is defined as

$$S = -\text{Tr}(\rho_i \log \rho_i). \quad (2.8)$$

Obviously, (2.2) is always satisfied by (2.5) with suitably chosen $\hat{\xi}_k$. But (2.5) does not satisfy (1.1), $[H, \rho_i] \neq 0$.

We are thus led to assume the density matrix in the form

$$\rho = \rho_i + \rho_1. \quad (2.9)$$

Actually, however, no approximation is introduced in so far as writing the density matrix in this form. For any given ρ , we can always find (2.5) with parameters ψ , $\hat{\xi}$, β and $\hat{\xi}_k$ so chosen that

$$\begin{aligned} \text{Tr}(\rho) &= \text{Tr}(\rho_i), \quad \text{Tr}(\rho N) = \text{Tr}(\rho_i N) \\ \text{Tr}(\rho H) &= \text{Tr}(\rho_i H), \quad \text{Tr}(\rho n_k) = \text{Tr}(\rho_i n_k). \end{aligned} \quad (2.10)$$

Then (2.9) can be regarded as the definition of ρ_1 , which should necessarily satisfy

$$\begin{aligned} \text{Tr}(\rho_1) &= 0, \quad \text{Tr}(\rho_1 N) = 0, \\ \text{Tr}(\rho_1 H) &= 0, \quad \text{Tr}(\rho_1 n_k) = 0. \end{aligned} \quad (2.11)$$

There are an infinite number of density matrices which satisfy (2.10). Among them, (2.5) is characterized by the maximum entropy and makes it possible to introduce thermodynamical variables as shown by (2.7). This is the reason for our particular choice (2.5), because the boundary condition of our problem is most simply expressed in terms of thermodynamical variables by specifying chemical potentials of sources at the boundary.

§ 3. Linear approximation

Now let us assume that the system is not far from equilibrium so that both $\hat{\epsilon}_k$ and ρ_1 are small quantities of first order. Then the well-known expansion formula of ordered exponential leads to

$$\rho_t = \rho_e \left(1 + \sum' \hat{\epsilon}_k \int_0^t \frac{d\lambda}{\beta} n_{-k}(-i\lambda) \right) \quad (3.1)$$

where we have made use of (2.4), i. e.,

$$\langle n_k \rangle_e = 0, \quad k \neq 0 \quad (3.2)$$

and $n_k(-i\lambda)$ is obtained from Heisenberg's operator $n_k(t) = \exp(iHt) \cdot n_k \cdot \exp(-iHt)$ by replacing t with $-i\lambda$.

Under the linear approximation (3.1), the average density is given as

$$\langle n_k \rangle = \text{Tr}(\rho_1 n_k) = g_k \hat{\epsilon}_k \quad (3.3)$$

where

$$g_k = \int_0^\beta \frac{d\lambda}{\beta} \langle n_{-k}(-i\lambda) n_k \rangle_e. \quad (3.4)$$

In particular, assuming g_k is continuous at $k=0$, we can show that

$$\lim_{k \rightarrow 0} g_k = \Omega (\partial n / \partial \hat{\epsilon}) \quad (3.5)$$

where n is defined by (2.4). From (2.7), the entropy is expressed as

$$S = S_0 - \frac{1}{2} \sum' g_k^{-1} \langle n_k \rangle \langle n_{-k} \rangle \quad (3.6)$$

with the equilibrium value S_0 . If the density fluctuation is gradual, i. e., only the $\hat{\epsilon}_k$ with small k are excited, we may replace g_k in (3.6) by (3.5), so that we obtain the well-known expansion of entropy in thermodynamics.

Let us now turn to the flow density which is defined by

$$\mathbf{v}(\mathbf{x}) = \frac{1}{2} \sum (\dot{\mathbf{x}}_j \delta(\mathbf{x} - \mathbf{x}_j) + \delta(\mathbf{x} - \mathbf{x}_j) \dot{\mathbf{x}}_j) = \Omega^{-1} \sum \mathbf{v}_k e^{i\mathbf{k} \cdot \mathbf{x}}. \quad (3.7)$$

In cases of neutral particles and also of charged particles without magnetic fields, the local equilibrium (3.1) makes no contribution to the average flow. In the case of charged particles moving in a magnetic field, on the contrary, the flow does not vanish even in

equilibrium, because there exists a diamagnetic current. This current, however, can be described in terms of a magnetization and, in particular, never gives rise to a net current flowing through a cross section of a conductor (except for superconductors). So we hereafter disregard this part of flow.

There still remains the flow due to the second term of (3.1)

$$\langle \mathbf{v}_k \rangle_t = \hat{\epsilon}_k \int_0^t \frac{d\lambda}{\beta} \langle n_{-k}(-i\lambda) \mathbf{v}_k \rangle_e. \quad (3.8)$$

This vanishes if there is no magnetic field. Because then n_k is invariant and \mathbf{r}_k changes its sign under the transformation of time reversal. In the presence of a magnetic field, however, the transformation of time reversal includes the reversal of the magnetic field, too. So we can infer only that (3.8) is reversed together with the magnetic field. In other words, the flow is antisymmetric with respect to the magnetic field. But it does vanish in the classical limit and therefore in the high temperature limit, too. In fact, in the classical limit, we can first perform the integration over particle velocities in taking the expectation value in (3.8) and this vanishes because it is linear in velocities.

A more detailed discussion of (3.8) will be given in § 5.

§ 4. Basic equation and its solution

Now inserting (2.9) together with (3.1) into (1.1), we obtain our basic equation

$$i[H, \rho_1] = - \sum_k \hat{\epsilon}_k \int_0^t \frac{d\lambda}{\beta} \rho_e \dot{n}_{-k}(-i\lambda) \quad (4.1)$$

where, of course,

$$\dot{n}_k \equiv i[H, n_k] = -ik \cdot \mathbf{v}_k. \quad (4.2)$$

A particular solution of the inhomogeneous equation (4.1) is given by

$$\rho_1 = \int_0^\infty dt e^{-\epsilon t} e^{-\epsilon H} I' e^{\epsilon H} \quad (4.3)$$

where I' stands for the right-hand side of (4.1). In fact, (4.3) is the solution which we are seeking for. This can be seen in the following way.

First, note that the macroscopic relaxation time, i.e., the relaxation time in which the system recovers equilibrium, is of the order of l^2/D . Here l is the linear dimension of the system and D the diffusion coefficient. Now suppose that the system was in equilibrium at the remote past, $t = -\infty$, and that the difference of chemical potentials at two ends of the system has been increased very slowly from zero to the present value ($t=0$). For instance, suppose that it is proportional to $\exp(\epsilon t)$, where $0 < \epsilon \ll D/l^2$. Then we may assume that a steady diffusion is established at each instant of time from $t = -\infty$ to $t=0$. This is a sort of adiabatic change and may be represented by the

solution proportional to $\exp(\varepsilon t)$ of the following Liouville equation

$$\begin{aligned}\dot{\rho}_1 + i[H, \rho_1] &= e^{\varepsilon t} I, \\ \rho_1(-\infty) &= 0, \quad \varepsilon \rightarrow +0.\end{aligned}$$

At $t=0$, we have

$$\varepsilon \rho_1 + i[H, \rho_1] = I \quad (4.4)$$

whose solution is nothing but (4.3), as can be easily confirmed.

It still remains to be confirmed that (4.3) does satisfy our subsidiary conditions (2.11). For any dynamical variable α , we have

$$\text{Tr}(\rho_1 \alpha) = - \sum' \hat{\varepsilon}_k \int_0^\infty dt e^{-\varepsilon t} \int_0^\beta \frac{d\lambda}{\beta} \langle \dot{n}_{-k}(-i\lambda) \alpha(t) \rangle_e. \quad (4.5)$$

If we take $\alpha=1$, the left-hand side is $\text{Tr}(\rho_1)$, and the right-hand side contains $\langle \dot{n}_{-k} \rangle_e$. This should vanish, because any flow in equilibrium is steady. The first condition in (2.11) is thus satisfied. As for the second condition, we take $\alpha=N$. Since N commutes with the Hamiltonian, the right-hand side of (4.5) contains $\langle \dot{n}_{-k} N \rangle_e$. This should also vanish, because

$$\text{Tr}(e^{\varepsilon N - \beta H} \dot{n}_{-k} N) = \frac{\partial}{\partial \varepsilon} \text{Tr}(e^{\varepsilon N - \beta H} \dot{n}_{-k}) = 0.$$

Thus the second condition is satisfied. Similarly, it can be seen by taking $\alpha=H$ that the third condition is also satisfied.

As for the final condition in (2.11), we have

$$\text{Tr}(\rho_1 n_k) = -i k \hat{\varepsilon}_k \cdot \int_0^\infty dt e^{-\varepsilon t} \int_0^\beta \frac{d\lambda}{\beta} \langle \mathbf{v}_{-k}(-i\lambda) n_k(t) \rangle_e \quad (4.6)$$

where we have made use of the continuity equation (4.2). In general, (4.6) does not vanish. We should, therefore, restrict ourselves to the case where the density fluctuation is so gradual that we may replace $\langle \dots \rangle_e$ in (4.6) by $\langle V N \rangle_e$. Here V is the net flow, i. e., \mathbf{v}_k with $k=0$. Since there is no net flow in equilibrium as mentioned before, we have

$$\text{Tr}(e^{\varepsilon N - \beta H} V N) = \frac{\partial}{\partial \varepsilon} \text{Tr}(e^{\varepsilon N - \beta H} V) = 0.$$

Thus, under our restriction, the fourth condition in (2.11) is satisfied.

The restriction can more precisely be expressed in the following way. In terms of space coordinates, (4.6) is written as

$$\text{Tr}(\rho_1 n(x)) = - \int dx' \mathbf{R}(x-x') \cdot \nabla \hat{\varepsilon}(x') \quad (4.7)$$

with

$$R(\mathbf{x}-\mathbf{x}') = \int_0^\infty dt e^{-\varepsilon t} \int_0^{\dot{\zeta}} \frac{d\lambda}{\beta} \langle \mathbf{v}(\mathbf{x}', -i\lambda) n(\mathbf{x}, t) \rangle_e.$$

Except for such singular systems as superconductors and superfluids, the relaxation function $R(\mathbf{x})$ will decay off within a finite distance. The gradient of chemical potential $\Gamma \dot{\zeta}$ should be nearly constant over this distance. This is the precise formulation of our restriction. If the condition is not satisfied, the simple thermodynamical notion of diffusion is not applicable in describing the phenomenon.

§ 5. Diffusion constant and Einstein's relation

Now, we take $\alpha=\mathbf{v}_k$ in (4.5), then the flow arising from ρ_1 is given as

$$\langle \mathbf{v}_k \rangle_1 \equiv \text{Tr}(\rho_1 \mathbf{v}_k) = -\dot{\zeta}_k \int_0^\infty dt e^{-\varepsilon t} \int_0^{\dot{\zeta}} \frac{d\lambda}{\beta} \langle \dot{n}_{-k}(-i\lambda) \mathbf{v}_k(t) \rangle_e. \tag{5.1}$$

Making use of the continuity equation (4.2), we see that in the limit of $k \rightarrow 0$ the diffusion equation takes the form

$$\begin{aligned} \langle v_{k\mu} \rangle_1 &= - \sum_\nu D_{\mu\nu}^{(1)} \times (ik_\nu \dot{\zeta}_k), \\ \mu, \nu &= x, y, z. \end{aligned} \tag{5.2}$$

In terms of space coordinates, this is written as

$$\langle v_\mu(\mathbf{x}) \rangle_1 = - \sum_\nu D_{\mu\nu}^{(1)} \frac{\partial \dot{\zeta}(\mathbf{x})}{\partial x_\nu}. \tag{5.3}$$

Here the diffusion tensor $D_{\nu\mu}$ is given by

$$D_{\mu\nu}^{(1)} = \int_0^\infty dt e^{-\varepsilon t} \int_0^{\dot{\zeta}} \frac{d\lambda}{\beta} \langle V_\nu(-i\lambda) V_\mu(t) \rangle_e \tag{5.4}$$

where V is the net flow.

On the other hand, it has been found¹⁾ that the electric conductivity tensor is given, in general, by

$$\sigma_{\mu\nu} = \int_0^\infty dt e^{-\varepsilon t} \int_0^{\dot{\zeta}} d\lambda \langle J_\nu(-i\lambda) J_\mu(t) \rangle_e \tag{5.5}$$

where $\mathbf{J}=eV$ is the electric current and e the charge of the particle. Comparing (5.4) with (5.5), we obtain the well-known Einstein relation

$$\sigma_{\mu\nu} = \beta e^2 D_{\mu\nu}^{(1)}. \tag{5.6}$$

It should be remembered, however, that in the presence of a magnetic field we have the flow arising from the local equilibrium distribution, (3.8), which is antisymmetric and thus makes contribution to the antisymmetric part of the diffusion tensor. Thus we

obtain the important conclusion that the Einstein relation is valid only for the symmetric part of diffusion and electric conductivity tensors.

In order to see the situation in more detail, let us take a system of electrons in a metal. For simplicity we shall neglect the interaction with phonons, taking account of impurity scattering alone. The Hamiltonian then takes the form

$$H = \sum_j H'(\mathbf{x}_j),$$

$$H'(\mathbf{x}) = \frac{1}{2m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A}(\mathbf{x}) \right)^2 + \varphi(\mathbf{x}),$$

where \mathbf{A} is the vector potential of the magnetic field and φ the scalar potential including both periodic and impurity potentials. Hereafter an operator of one electron will be primed as H' . Let us introduce the system of one electron eigenfunctions

$$H' u_r = E_r u_r.$$

Then the Hamiltonian can be written as

$$H = \sum_r E_r a_r^* a_r, \quad (5.7)$$

where a_r , a_r^* are destruction and creation operators of electrons in the r -state. Similarly

$$n_k = \sum \langle r | n_k' | s \rangle a_r^* a_s, \quad (5.8)$$

$$\mathbf{v}_k = \sum \langle r | \mathbf{v}_k' | s \rangle a_r^* a_s,$$

where

$$n_k' = e^{ik \cdot x}, \quad \mathbf{v}_k' = \frac{1}{2} (\dot{\mathbf{x}} e^{ik \cdot x} + e^{ik \cdot x} \dot{\mathbf{x}}),$$

$$\dot{\mathbf{x}} = i[H', \mathbf{x}]. \quad (5.9)$$

Remember that

$$e^{iHt} a_r^* a_s e^{-iHt} = a_r^* a_s e^{i(E_r - E_s)t}$$

and also that

$$\langle a_s^* a_r a_r^* a_s \rangle_e = \delta_{rs} \delta_{rr'} f(E_s) (1 - f(E_r))$$

where f is the Fermi distribution function

$$f(E) = [\exp(\beta E - \hat{\epsilon}) + 1]^{-1}. \quad (5.10)$$

Then it is easily seen that (5.1) can be written as

$$\langle \mathbf{v}_k \rangle_1 = -\hat{\epsilon}_k \sum_l \frac{f(E_r) - f(E_s)}{\beta(E_r - E_s)} \langle s | n_{-k}' | r \rangle \langle r | \mathbf{v}_k' | s \rangle \left\{ \pi \delta(E_r - E_s) + \frac{P}{i(E_r - E_s)} \right\} \quad (5.11)$$

where P indicates taking the principal value. The flow arising from this part is anti-symmetric, whereas the flow arising from the δ -function is symmetric.

Similarly, (3.8) can be written as

$$\langle \mathbf{v}_k \rangle_l = \hat{\epsilon}_k \sum_l \frac{f(E_r) - f(E_s)}{\beta(E_r - E_s)} \langle s | n_{-k}' | r \rangle \langle r | \mathbf{v}_k' | s \rangle. \quad (5.12)$$

For the sum of those terms in which $E_r \approx E_s$, we can make use of

$$\langle s | \dot{n}'_{-k} | r \rangle = i(E_s - E_r) \langle s | n'_{-k} | r \rangle \quad (5.13)$$

so that the sum exactly cancels the antisymmetric part of (5.11). On the other hand, the terms with $E_r = E_s$ in (5.12) can be transformed as

$$\begin{aligned} \hat{\epsilon}_k \sum_s \frac{1}{\beta} \frac{\partial f(E_s)}{\partial E} \langle s | \frac{1}{2} (n'_{-k} \mathbf{v}_k' + \mathbf{v}_k' n'_{-k}) | s \rangle - \hat{\epsilon}_k \sum_{E_r \neq E_s} \sum_{E_s} \frac{1}{\beta} \frac{\partial f(E_s)}{\partial E} \\ \times \frac{1}{2} \{ \langle s | n'_{-k} | r \rangle \langle r | \mathbf{v}_k' | s \rangle + \langle s | \mathbf{v}_k' | r \rangle \langle r | n'_{-k} | s \rangle \}. \end{aligned} \quad (5.14)$$

From (5.9) we see

$$\frac{1}{2} (n'_{-k} \mathbf{v}_k' + \mathbf{v}_k' n'_{-k}) = \dot{\mathbf{x}}$$

so that the first term of (5.14) vanishes. For the second term, we can again make use of (5.13). Taking the limit of $k \rightarrow 0$ we find the antisymmetric part of the diffusion tensor

$$D_{\mu\nu}^{(A)} = \sum_s \frac{1}{\beta} \frac{\partial f(E_s)}{\partial E} \cdot \frac{P}{i(E_s - E_r)} \times \frac{1}{2} \{ \langle s | \dot{x}_\nu | r \rangle \langle r | \dot{x}_\mu | s \rangle - \langle s | \dot{x}_\mu | r \rangle \langle r | \dot{x}_\nu | s \rangle \}. \quad (5.15)$$

On the other hand, the antisymmetric part of the electric conductivity tensor is given as

$$\frac{1}{\beta e^2} \sigma_{\mu\nu}^{(A)} = \sum_s \frac{f(E_r) - f(E_s)}{\beta(E_r - E_s)} \cdot \frac{P}{i(E_s - E_r)} \times \frac{1}{2} \{ \langle s | \dot{x}_\nu | r \rangle \langle r | \dot{x}_\mu | s \rangle - \langle s | \dot{x}_\mu | r \rangle \langle r | \dot{x}_\nu | s \rangle \}. \quad (5.16)$$

Now, characteristic frequencies of the electron velocity $\langle s | \dot{x} | r \rangle \exp\{i(E_r - E_s)t\}$ are cyclotron frequency ω_c , collision frequency τ^{-1} , and also the average interval of interband transitions ΔE . If all these frequencies satisfy

$$\beta\omega_c \ll 1, \quad \beta\tau^{-1} \ll 1, \quad \beta\Delta E \ll 1 \quad (5.17)$$

then the difference quotient of f in (5.16) is practically the same as the differential quotient in (5.15), and we have the Einstein relation for the antisymmetric part, too. It is to be noted here that the second of (5.17) is the well-known criterion for applicability of Bloch's kinetic equation, although actually this is too stringent.

§ 6. Damping theoretical expansion

Finally, we shall show that our basic equation (4.1) reduces to the usual Bloch equation in the limit of a system of weakly interacting particles. In fact (4.1) has mathematically the same form as discussed by Kohn and Luttinger¹⁰ in the case of a steady electric field. They have dealt with a simple model in which an electron is scattered by impurities. Actually their method can be applied to integration of the Liouville equation in general, provided that the Hamiltonian satisfies certain conditions.

First let us introduce a symbolic method due to Kubo. Define the linear operator \mathcal{L} which operates always from the left on any dynamical variable η as

$$\mathcal{L} \cdot \eta \equiv i[H, \eta]. \quad (6.1)$$

Incidentally, the operator of this sort has been introduced in classical statistical mechanics to prove the ergodic theorem.⁷⁾ The Liouville equation can be written as

$$\left(\frac{\partial}{\partial t} + \mathcal{L} \right) \rho(t) = 0. \quad (6.2)$$

Let us introduce the Laplace transform

$$\rho(s) = \int_0^\infty dt e^{-st} \rho(t) \quad (6.3)$$

which converges for $\Re s > 0$. Then (6.2) is transformed into

$$(s + \mathcal{L}) \rho(s) = \rho_0 \quad (6.4)$$

where ρ_0 is the initial density matrix at $t=0$. Replacing $\rho(s)$ by ρ_1 , ρ_0 by Γ , and taking the limit $s \rightarrow +0$, we obtain our basic equation (4.4). But we shall be concerned with the general case (6.4) for the moment.

Now, assume that the Hamiltonian takes the form

$$H = H_0 + gH_I \quad (6.5)$$

where g is a small numerical parameter indicating the order of perturbation H_I . Correspondingly the operator \mathcal{L} splits as

$$\mathcal{L} = \mathcal{L}_0 + g\mathcal{L}_I. \quad (6.6)$$

Let us fix the basic state vectors $|\alpha\rangle$ as

$$H_0|\alpha\rangle = \omega_\alpha|\alpha\rangle.$$

In the case of the electron-phonon interaction, for instance, α stands for a set of occupation numbers of free electrons and phonons. We introduce linear operators \mathcal{D} and \mathcal{H} which also operate from the left on any dynamical variable as

$$\begin{aligned} \langle \alpha | \mathcal{D} \eta | \alpha' \rangle &= \langle \alpha | \eta | \alpha' \rangle \delta_{\alpha\alpha'}, \\ \langle \alpha | \mathcal{H}(\eta) | \alpha' \rangle &= \langle \alpha | \eta | \alpha' \rangle (1 - \delta_{\alpha\alpha'}). \end{aligned}$$

Without loss of generality we assume that $\mathcal{L}H_I = 0$. Following Kohn and Luttinger, we decompose the density matrix into diagonal and non-diagonal parts

$$\begin{aligned} \rho(s) &= \rho_d(s) + \rho_n(s), \\ \rho_d &= \mathcal{D}\rho, \quad \rho_n = \mathcal{H}\rho. \end{aligned} \quad (6.7)$$

Inserting (6.7) together with (6.6) into (6.4) and taking diagonal and non-diagonal parts of the equation respectively, we obtain

$$\begin{aligned} s\rho_d(s) + g\mathcal{D}\mathcal{L}_I\rho_n(s) &= \mathcal{D}\rho_0, \\ (s + \mathcal{L}_0 + g\mathcal{H}(\mathcal{L}_I))\rho_n(s) &= \mathcal{H}(\rho_0 - g\mathcal{L}_I\rho_d(s)). \end{aligned} \quad (6.8)$$

Eliminating ρ_n from these, we have a formal solution for the diagonal part

$$s\rho_d(s) - g^2 \mathcal{D} \mathcal{L}_1 \frac{1}{s + \mathcal{L}_0 + g\mathcal{H}(\mathcal{L}_1)} \mathcal{L}_1 \rho_d(s) = \mathcal{D} \left(\rho_0 - \frac{1}{s + \mathcal{L}_0 + g\mathcal{H}(\mathcal{L}_1)} \mathcal{H}(\rho_0) \right). \quad (6.9)$$

We can derive a power series expansion from (6.9), supposing that

$$\rho_d(s) = O(g^{-2}) \quad (6.10)$$

and making use of

$$\frac{1}{s + \mathcal{L}_0 + g\mathcal{H}(\mathcal{L}_1)} = \frac{1}{s + \mathcal{L}_0} + \frac{1}{s + \mathcal{L}_0} g\mathcal{H}(\mathcal{L}_1) \frac{1}{s + \mathcal{L}_0} + \dots \quad (6.11)$$

In particular the lowest order approximation satisfies

$$s\rho_d^{(0)}(s) - g^2 \mathcal{D} \mathcal{L}_1 \frac{1}{s + \mathcal{L}_0} \mathcal{L}_1 \rho_d^{(0)}(s) = \mathcal{D} \rho_0. \quad (6.12)$$

Remember that

$$\frac{1}{s + \mathcal{L}_0} \eta = \int_0^\infty dt \exp(-st) \cdot \exp(-iH_0 t) \cdot \eta \exp(iH_0 t).$$

Then, taking the explicit matrix representation of (6.12), we find

$$sP_\alpha(s) - g^2 \sum_{\alpha'} \frac{2s}{s^2 + (\omega_\alpha - \omega_{\alpha'})^2} |\langle \alpha | H_1 | \alpha' \rangle|^2 (P_{\alpha'}(s) - P_\alpha(s)) = \langle \alpha | \rho_0 | \alpha \rangle, \quad (6.13)$$

$$P_\alpha(s) = \langle \alpha | \rho_d^{(0)}(s) | \alpha \rangle.$$

In the limit of $s \rightarrow +0$, we have

$$\frac{2s}{s^2 + (\omega_\alpha - \omega_{\alpha'})^2} \sim 2\pi \delta(\omega_\alpha - \omega_{\alpha'}).$$

Hence (6.13) is nothing else but the Laplace transform of the so-called master equation

$$\begin{aligned} \dot{P}_\alpha(t) &= \sum 2\pi g^2 |\langle \alpha | H_1 | \alpha' \rangle|^2 \delta(\omega_\alpha - \omega_{\alpha'}) (P_{\alpha'}(t) - P_\alpha(t)), \\ P_\alpha(t=0) &= \langle \alpha | \rho_0 | \alpha \rangle. \end{aligned} \quad (6.14)$$

which van Hove²⁾ has obtained by means of rather a lengthy expansion.

Now replacing ρ by ρ_1 , ρ_0 by I , and taking the limit $s \rightarrow +0$, we have

$$2\pi g^2 \sum_{\alpha'} |\langle \alpha | H_1 | \alpha' \rangle|^2 \delta(\omega_\alpha - \omega_{\alpha'}) (F_{\alpha'} - F_\alpha) = \langle \alpha | I | \alpha' \rangle \quad (6.15)$$

where F_α is the diagonal element of ρ_1 in the lowest order approximation. The equation (6.15) says that the drift term due to the gradient of chemical potential is balanced by the collision term due to perturbation H_1 . More precisely, F_α is still a many-particle distribution function and we have further to reduce (6.15) to obtain the Bloch equation of one particle distribution function.

Of course, certain conditions should be satisfied in order that the transport equation

(6·12) be already a good approximation. Smallness of the parameter g is by no means sufficient. For instance, in the case of an electron scattered by impurities, the contribution of the second term in (6·11) relative to the first would not be simply proportional to g , but to $g\Omega$, if impurities were on regular lattice points. In this case the transport equation can never be a good approximation, because $g\Omega$ can be increased indefinitely by increasing the volume Ω . In general, the size dependence of perturbation in a large system plays an essential role here. We shall not enter into this problem, as the detailed analyses have been given by van Hove.⁹⁾

§ 7. Conclusions

Quantum statistical mechanics of steady diffusion has been formulated. The theory is the most natural generalization of the Bloch kinetic equation and in fact reduces to the latter under certain conditions. An important conclusion is that the Einstein relation is valid only for the antisymmetric part of diffusion- and electric conductivity tensors.

Obviously, the theory can be generalized so as to include heat conduction and viscosity, which will be discussed in a subsequent paper.

The expressions for transport coefficients derived in this way are general and rigorous. In statistical thermodynamics, we have the general expression for the partition function and introduce approximate methods in evaluating this; the virial expansion in the case of imperfect gases, normal vibrations in the case of crystals, and so on. In just the same way, we should introduce appropriate methods of approximation to evaluate transport coefficients, starting with our general expressions. Thus we may conclude that quantum theory of transport coefficients now stands on the same level as statistical mechanics of equilibrium properties, though admittedly we know few methods of approximation such as Bloch's kinetic equation which is nothing but the lowest order approximation in a damping theoretical expansion.

References

- 1) H. Nakano, *Prog. Theor. Phys.* **15** (1956), 77.
R. Kubo, *J. Phys. Soc. Japan* **12** (1957), 570.
- 2) R. Kubo, M. Yokota, and S. Nakajima, *J. Phys. Soc. Japan* **12** (1957), 1203.
- 3) W. Kohn and J. M. Luttinger, *Phys. Rev.* **108** (1957), 590.
- 4) H. B. Callen, M. L. Barasch and J. L. Jackson, *Phys. Rev.* **88** (1952), 1382.
- 5) R. Kubo, N. Hashitsume and M. Yokota, *Busseiron Kenkyu*, **91** (1955), 32.
- 5) S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases*, Cambridge, 1939.
- 6) T. Matsubara, *Busseiron Kenkyu*, vol. 2 (1957), 190.
- 7) E. Hopf, *Ergodentheorie*, Berlin, 1937.
- 8) L. van Hove, *Physica* **21** (1955), 517; *ibid.* **23** (1957), 441.

Foundation of Deformed Potential Model for Nuclear Rotation

Kiyomi IKEDA, Minoru KOBAYASI, Toshio MARUMORI,*
Sinobu NAGATA and Takanori SHIOZAKI

Department of Physics, Kyoto University, Kyoto

(Received August 15, 1958)

Following the prescription suggested by Bohr and Mottelson, we develop a formulation of the separation of the nuclear Hamiltonian into rotational and intrinsic parts. In the light of this formulation, the foundation of various methods with the deformed potential model for the estimation of the moments of inertia for nuclear rotation is discussed and the interrelation among these methods is clarified. Treatments are restricted to the two dimensional case for simplicity.

§ 1. Introduction

The fact that the existence of rotational spectra had been confirmed in certain regions of elements brought forward the problem of finding the transformation which leads, at least approximately, to the separation of the nuclear Hamiltonian into intrinsic and rotational parts.

The early workers¹ in this field succeeded in indicating the existence of rotational states of nuclei qualitatively, but they could say nothing quantitative about the values of the moments of inertia associated with these rotational motions. They gave much smaller values for the moments of inertia than those empirically determined. It has become clear in time this discrepancy originates in the poor separation of the nuclear Hamiltonian.

On the other hand, various prescriptions for the calculation of the moment of inertia without recourse to the coordinate transformation were also proposed by many authors.²⁻⁵ Among those, the so-called generator coordinates method^{2,3} and the cranking model^{4,5} are most typical. The common characteristic of these methods is that the wave functions corresponding to the independent particle motions in a deformed Hartree field are used as the model wave functions to evaluate the moment of inertia. Hence we may call them the methods with the deformed potential model.

Many authors derived different expressions²⁻⁵ for the moment of inertia of the rotating nucleus by making use of this deformed potential model. But, it was difficult to discuss the relative merits of these results, because the physical foundations of such calculations were not so clear. Further, as was pointed out by Lipkin,⁶ the question that at what stage of the calculation the model wave function is to be adopted has been left open.

Thus, there are two problems in our face: the first problem is to find the compact

* The author thanks the Iwanami Fujukai for the financial aid.

form of transformation which leads to the optimum separation of the nuclear Hamiltonian into rotational and intrinsic parts; the second problem is to show how and at what stage of transformation the appropriate model wave functions should be applied in the calculation of the moment of inertia of the rotating nucleus.

As for the first problem, Bohr and Mottelson¹⁰⁾ have recently suggested a very ingenious prescription for the separation of the nuclear Hamiltonian by introducing a kind of canonical transformation which adopts collective orientation angles depending on both the positions and the momenta of the nuclear particles. Indeed, the ordinary point transformations leave so large an amount of coupling terms that they do not even approximately separate the Hamiltonian into two parts.⁷⁾¹⁰⁾⁻¹⁴⁾ Further, these authors gave a condition for separation of the nuclear Hamiltonian in the form of operator equations, though the satisfactory test for the applicability of this condition in the case of real nuclei is impossible at the present stage of the theory. This problem seems to be closely connected with the nature of nuclear forces. That is, the criterion to the separation of the nuclear Hamiltonian depends essentially on the interactions between nucleons.

With regard to the second problem, it must be emphasized that the model Hamiltonian with the deformed potential is not invariant under rotation, so that the model wave functions cannot be the eigen-functions of the total angular momentum. This situation causes some confusions in the calculation of the nuclear moment of inertia. The moment of inertia calculated by this model Hamiltonian does not correspond to that for nuclear free rotation, but to that for the spurious oscillation about the equilibrium orientation in the deformed potential fixed in the space⁹⁾.

In this paper, we shall first, in § 2, develop the method of canonical transformation introduced by Bohr and Mottelson¹⁰⁾ in the formalism of the redundant variables with the subsidiary condition and show on what condition the separation of the real nuclear Hamiltonian is performed. Next, in § 3, the foundation and the validity of various methods with the deformed potential model will be discussed from the standpoint of the canonical transformation formalism developed in § 2. The question raised by Lipkin⁹⁾ will be answered by showing the step of the transformation at which the model wave function should be applied. Finally, in § 4, we shall discuss the validity of the various expressions for the moment of inertia of the rotating nucleus obtained by many authors.

§ 2. Separation of the nuclear Hamiltonian into rotational and intrinsic parts

Let us consider such an even-even nucleus but not of a closed-shell configuration that exhibits rotational spectra. The real Hamiltonian H_0 is

$$H_0 = \sum_{i=1}^A T_i + \sum_{i < j} v_{ij} \quad (2 \cdot 1)$$

which is invariant under the spatial rotation. Here T_i is the kinetic energy of the i -th nucleon and v_{ij} is the interaction potential between the i -th and the j -th nucleons.

For simplicity we restrict ourselves to a two dimensional problem. Since we are

interested in the rotational spectra of the form

$$E_{i,I} = \epsilon_i + \frac{1}{2\mathcal{I}} I^2, \quad (2.2)$$

our Schrödinger equation may be given in the form

$$H_0 \Psi_{iI}(\mathbf{x}_1, \dots, \mathbf{x}_A) = E_{iI} \Psi_{iI}(\mathbf{x}_1, \dots, \mathbf{x}_A), \quad (2.3)$$

where Ψ_{iI} is the eigenstate belonging to a particular rotational band ϵ_i and having an eigenvalue I of the total angular momentum \mathbf{I} .

We shall adopt the method of the canonical transformation with the redundant variables by introducing redundant variables θ and \mathbf{I}_0 which satisfy the commutation relation.

$$[\mathbf{I}_0, \theta] = -i\hbar. \quad (2.4)$$

Then, the Schrödinger equation (2.3) becomes

$$H_0 \Psi_{iI}^{(0)}(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta) = E_{iI} \Psi_{iI}^{(0)}(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta) \quad (2.5a)$$

with the subsidiary condition,

$$\mathbf{I}_0 \Psi_{iI}^{(0)}(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta) = 0. \quad (2.5b)$$

This subsidiary condition is formally equivalent to the condition

$$\theta \Psi_{iI}^{(0)}(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta) = 0,$$

which is commonly used in the redundant variable method, but our condition (2.5b) is more physical than the usual one and is easily extended to the three dimensional case.*

Next, we perform the following canonical transformations successively,

$$\Psi_{iI}^{(3)} = (U_1 U_2 U_3)^{-1} \Psi_{iI}^{(0)}, \quad (2.6)$$

$$U_1 = \exp\left(-\frac{i}{\hbar} \theta \mathbf{I}_0\right), \quad (2.7)$$

$$U_2 = \exp\left(+\frac{i}{\hbar} \theta \mathbf{I}\right), \quad (2.8)$$

$$U_3 = \exp\left(-\frac{i}{\hbar} \theta \mathbf{I}_0\right), \quad (2.9)$$

where \mathbf{I} is the angular momentum of all particles in terms of the particle coordinates

$$\mathbf{I} = \sum_i (\mathbf{x}_i \mathbf{p}_i^t - \mathbf{y}_i \mathbf{p}_i^x),$$

and θ is a collective orientation angle, which is generally a function of both positions and momenta of the particles.

As was shown by Bohr and Mottelson,¹⁰ if \mathbf{I} and θ satisfy the following conditions,

$$\frac{i}{\hbar} [H_0, \theta] = \frac{1}{2\mathcal{I}} \mathbf{I}, \quad (2.10)$$

* On this point, see reference 14).

$$[\mathbf{I}, \mathcal{J}] = -i\hbar, \quad (2.11)$$

the transformed Hamiltonian and the subsidiary condition are

$$H_3 \equiv (U_1 U_2 U_3)^{-1} H_0 U_1 U_2 U_3 = \frac{1}{2\mathcal{J}} \mathbf{I}_0^2 + \left(H_0 - \frac{1}{2\mathcal{J}} \mathbf{I}^2 \right), \quad (2.12a)$$

$$\mathbf{I} \Psi_s^{(3)}(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta) = 0. \quad (2.12b)$$

This representation is the so-called collective representation, where rotational and intrinsic parts are completely separated. In fact, Hamiltonian H_3 contains neither \mathcal{J} nor \mathbf{I} , that is, it satisfies

$$[H_3, \mathcal{J}] = [H_3, \mathbf{I}] = 0. \quad (2.13)$$

This implies that we have displayed the dependence of the Hamiltonian (2.1) on the total angular momentum without making explicit use of the intrinsic variables which should be complicated functions of particle variables $\mathbf{x}_i, \mathbf{p}_i$.*

In this collective representation, $\Psi_{iI}^{(3)}$ is written in the following form,

$$\Psi_{iI}^{(3)}(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta) = \chi_I(\theta) \varphi_i(\mathbf{x}_1, \dots, \mathbf{x}_A). \quad (2.14)$$

Here the intrinsic wave function $\varphi_i(\mathbf{x}_1, \dots, \mathbf{x}_A)$ and the rotational wave function $\chi_I(\theta)$ satisfy the following equations respectively,

$$H_{\text{in}} \varphi_i \equiv \left(H_0 - \frac{1}{2\mathcal{J}} \mathbf{I}^2 \right) \varphi_i(\mathbf{x}_1, \dots, \mathbf{x}_A) = \epsilon_i \varphi_i(\mathbf{x}_1, \dots, \mathbf{x}_A), \quad (2.15a)$$

$$\mathbf{I} \varphi_i(\mathbf{x}_1, \dots, \mathbf{x}_A) = 0, \quad (2.15b)$$

$$\frac{1}{2\mathcal{J}} \mathbf{I}_0^2 \chi_I(\theta) = \frac{1}{2\mathcal{J}} I^2 \chi_I(\theta), \quad (2.16)$$

and the energy of the system becomes

$$E_{iI} = \epsilon_i + \frac{1}{2\mathcal{J}} I^2, \quad (2.2)$$

as was expected in the beginning of this section.

If we use the Hartree approximation in (2.15a) at this stage, the wave function φ_i in (2.15a) which describes the intrinsic motion is replaced by the model wave function describing the independent particle motion in the deformed field. Then, in general, this model wave function cannot satisfy the subsidiary condition (2.15b). On the other hand, it is inevitable to use the model wave functions in order to give the explicit expression for the moment of inertia. In fact, if we adopt the correct wave function φ_i , the expressions for the moment of inertia, obtained by Inglis⁴⁾ and Skyrme⁵⁾ respectively,

* In this representation the total angular momentum becomes

$$-\mathbf{I}_0 = (U_1 U_2 U_3)^{-1} \mathbf{I} (U_1 U_2 U_3).$$

$$\mathcal{J} = 2 \sum_{j \neq i} \frac{\langle j | \mathbf{I} | i \rangle^2}{\epsilon_j - \epsilon_i},$$

$$\mathcal{J} = \frac{1}{2} \frac{\langle i | \mathbf{I}^2 | i \rangle - \langle i | \mathbf{I}^2 | i \rangle^2}{\langle i | H_0 \mathbf{I}^2 | i \rangle - \langle i | H_0 | i \rangle \langle i | \mathbf{I}^2 | i \rangle},$$

have no significance because of (2.15b). Here it is noted that the energy-eigenvalues ϵ_i used in the above expressions are those of H_1 . Consequently, such an introduction of the model wave function at this stage seems to contain a serious inconsistency in itself. It seems to us that these circumstances caused some confusions in the previous calculations of the moment of inertia. In § 3, we shall discuss this problem in detail.

§ 3. Introduction of the model wave function

In order to search for the representation which makes possible to introduce the model wave function, let us return to the representation of the second step,

$$\mathcal{F}_{iI}^{(2)} = (U_1 U_2)^{-1} \mathcal{F}_{iI}^{(0)}. \quad (3.1)$$

In this representation the Hamiltonian H_2 takes the same form as H_3 ,

$$H_2 \equiv (U_1 U_2)^{-1} H_0 U_1 U_2 = \frac{1}{2\mathcal{J}} \mathbf{I}_0^2 + \left(H_0 - \frac{1}{2\mathcal{J}} \mathbf{I}^2 \right) \quad (3.2)$$

but the subsidiary condition is

$$(\mathbf{I}_0 + \mathbf{I}) \mathcal{F}_{iI}^{(2)} = 0. \quad (3.3)$$

Let us further make a time dependent unitary transformation,

$$\phi_{iI} = U^{-1} \mathcal{F}_{iI}^{(2)} \quad (3.4)$$

$$U = \exp \left(- \frac{i}{\hbar} \frac{\mathbf{I}_0^2}{2\mathcal{J}} t \right). \quad (3.5)$$

The wave function ϕ_{iI} satisfies the following equations

$$i\hbar \frac{\partial \phi_{iI}}{\partial t} = \left(H_0 - \frac{1}{2\mathcal{J}} \mathbf{I}^2 \right) \phi_{iI}, \quad (3.6a)$$

$$(\mathbf{I}_0 + \mathbf{I}) \phi_{iI} = 0, \quad (3.6b)$$

and the equation of motion of $\theta(t)$ is

$$\frac{d\theta(t)}{dt} \equiv \omega = \frac{i}{\hbar} \left[\frac{1}{2\mathcal{J}} \mathbf{I}_0^2, \theta(t) \right] = \frac{1}{\mathcal{J}} \mathbf{I}_0. \quad (3.7)$$

The operator $H_0 - \frac{1}{2\mathcal{J}} \mathbf{I}^2$ in (3.6a) is time-independent, so that (3.6) satisfies the following eigenvalue equation,

$$\left(H_0 - \frac{1}{2\mathcal{J}} \mathbf{I}^2 \right) \phi_{iI} = \epsilon_i \phi_{iI} \quad (3.8a)$$

$$(\mathbf{I}_0 + \mathbf{I})\Phi_{iI} = 0. \quad (3.8b)$$

Here it is clear that the eigenvalue ϵ_i has the same value as that of ϵ_i in (2.2). Taking into account the subsidiary condition (3.8b), the equations (3.8) can be written in the form,

$$\left(H_0 - \frac{1}{2} \mathbf{I}_0^2\right) \Phi_{iI}(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta) = \epsilon_i \Phi_{iI}(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta), \quad (3.9a)$$

$$(\mathbf{I}_0 + \mathbf{I}) \Phi_{iI} = 0. \quad (3.9b)$$

It should be noted that Φ_{iI} is the eigenstate of the total angular momentum.

On the other hand, however, the wave functions

$$\phi_i = \sum_I a_I \Psi_{iI}(\mathbf{x}_1, \dots, \mathbf{x}_A) \chi_I^*(\theta) \quad (3.10)^*$$

generally satisfy also the same equations as (3.9),

$$\left(H_0 - \frac{1}{2} \mathbf{I}_0^2\right) \phi_i = \epsilon_i \phi_i, \quad (3.11a)$$

$$(\mathbf{I}_0 + \mathbf{I}) \phi_i = 0. \quad (3.11b)$$

Here $\chi_I(\theta)$ and Ψ_{iI} are defined through (2.16) and (2.3), respectively, and a_I 's are arbitrary coefficients of superposition. This wave function ϕ_i is not the eigenfunction of the total angular momentum. In fact, if we transform ϕ_i back into the original representation, it will be confirmed that the transformed ϕ_i is not an eigenstate of H_0 , but a superposition of the eigenstates of H_0 . In spite of this property, ϕ_i satisfies the subsidiary condition (3.11b). This is due to the particular choice of the present representation.

This fact assures the possibility of replacing the actual wave function Φ_{iI} by the model wave function ϕ_i which is not the eigenfunction of the total angular momentum, so long as we calculate the intrinsic energy ϵ_i characterizing the rotational band. Since the coefficient a_I in ϕ_i is not restricted, it might be possible to replace ϕ_i by the model wave function ϕ^M used in the deformed potential model. Anyhow the model wave function ϕ^M in the deformed potential model should be introduced only in this representation.

From this point of view, let us consider the validity of the model wave function ϕ^M used in the deformed potential model. The model wave function $\phi_i^M(\mathbf{x}, \theta)$ may be introduced as follows. Remark that

$$\phi_i(\mathbf{x}, \theta) = \exp\left(-\frac{i}{\hbar} \mathbf{I}\theta\right) \phi_i(\mathbf{x}, 0), \quad (3.12a)^{**}$$

$$\phi_i(\mathbf{x}, 0) = \sum_I a_I \Psi_{iI}(\mathbf{x}_1, \dots, \mathbf{x}_A), \quad (3.12b)$$

* Here $\chi_I^*(\theta)$ is complex conjugate to $\chi_I(\theta)$.

** \mathbf{x} in $\phi_i(\mathbf{x}, \theta)$ stands for a set of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_A$.

and $\phi_i(\mathbf{x}, 0)$ satisfies the following equations,

$$\left(H_0 - \frac{1}{2\mathcal{J}}\mathbf{I}^2\right)\phi_i(\mathbf{x}, 0) = \epsilon_i \phi_i(\mathbf{x}, 0), \quad (3.13a)$$

$$\mathbf{I}_0 \phi_i(\mathbf{x}, 0) = 0. \quad (3.13b)$$

Here, when the model wave function $\phi_i^M(\mathbf{x}, 0)$ in the deformed potential model is adopted in place of $\phi_i(\mathbf{x}, 0)$, $\phi_i^M(\mathbf{x}, \theta)$ is given as

$$\phi_i^M(\mathbf{x}, \theta) = \exp\left(-\frac{i}{\hbar}\mathbf{I}\theta\right)\phi_i^M(\mathbf{x}, 0). \quad (3.14)$$

Now, if we expand $\phi_i^M(\mathbf{x}, 0)$ in terms of the complete set $\mathcal{V}_k(\mathbf{x}_1, \dots, \mathbf{x}_A)$ which are the eigenstates of H_0 and \mathbf{I} , $\phi_i^M(\mathbf{x}, 0)$ may be expressed generally as follows.

$$\phi_i^M(\mathbf{x}, 0) = \sum_k a_k \mathcal{V}_k(\mathbf{x}_1, \dots, \mathbf{x}_A) \quad (3.15)$$

$$= \sum_j b_j \phi_j(\mathbf{x}, 0) + \text{terms not belonging to the rotational states.}$$

Therefore,

$$\phi_i^M(\mathbf{x}, \theta) = b_i \phi_i(\mathbf{x}, \theta) + \sum_{j \neq i} b_j \phi_j(\mathbf{x}, \theta) + \text{terms not belonging to the rotational states.} \quad (3.16)$$

Here, we have taken into account that $\phi_i(\mathbf{x}, \theta)$ is a superposition only of the rotational states belonging to a particular band ϵ_i in all eigenstates of H_0 .

The second and third terms in the right hand side of (3.16), which are deviations from (3.10), decide the validity of the adopted model wave function $\phi_i^M(\mathbf{x}, \theta)$. These terms have caused the serious confusion on the calculation of the moment of inertia. We shall discuss this point in detail in § 4.

On the other hand, the above formulation gives new light upon the various methods with the deformed potential. At first, it is to be noted that the equation (3.11a) is just the same as those satisfied by the extended wave function $F(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta)$ in Skyrme's method⁸. This equation is the starting point of his variational method. In his method, emphasis is placed on a band of rotational states rather than a single state and there the additional collective coordinate θ is not redundant.

Next, if we multiply the equation (3.10) by $\chi_I(\theta)$ and integrate over θ , we obtain

$$\mathcal{V}_{iI}(\mathbf{x}_1, \dots, \mathbf{x}_A) = \frac{1}{a_I} \int \chi_I(\theta) \phi_i(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta) d\theta. \quad (3.17)$$

This expression of \mathcal{V}_{iI} is simply that given by Peierls and Yoccoz's generator coordinate method²⁾ when $\phi_i(\mathbf{x}_1, \dots, \mathbf{x}_A, \theta)$ is replaced by the $\phi_i^M(\mathbf{x}, \theta)$. However, it should not be confused that ϕ_i^M does not correspond to the Hartree approximation of \mathcal{V}_{iI} , but to that of the intrinsic wave function ϕ_{iI} .

Thirdly, let us consider the cranking model proposed by Inglis and others.^{4, -7)} In this model the rotational motion is treated in a classical way. In our case this picture

corresponds to (3.6) and (3.7), where the rotational motion is described in the Heisenberg picture. In the cranking model, emphasis is placed on the motion of $\theta(t)$ and on setting up a deformed potential explicitly as an operator. Therefore, the equation (3.6) is written as follows,

$$i\hbar \frac{\partial \phi^{CM}}{\partial t} = \left\{ \sum_{i=1}^A T_i + \sum_{i=1}^A V_i(r_i, \varphi_i - \theta(t)) \right\} \phi^{CM} = H^{CM} \phi^{CM}, \quad (3.18a)$$

$$(\mathbf{I}_0 + \mathbf{I}) \phi^{CM} = 0. \quad (3.18b)$$

The angular dependence of V is required to satisfy the condition,

$$[H^{CM}, \mathbf{I}_0 + \mathbf{I}] = 0. \quad (3.19)$$

In this case, it should be kept in mind that the Hamiltonian H^{CM} is time-dependent. The equation (3.18a) is just the starting point of the cranking model.

In the cranking model the calculation of the moments of inertia is performed on the *body-fixed rotating coordinate system*. This corresponds to the transformation,

$$\phi'^{CM} = R(t)^{-1} \phi^{CM} \quad (3.20)$$

$$R(t) = \exp\left(-\frac{i}{\hbar} \mathbf{I} \theta(t)\right). \quad (3.21)$$

Then (3.18) become

$$i\hbar \frac{\partial \phi'^{CM}}{\partial t} = \left\{ \sum_{i=1}^A T_i + \sum_{i=1}^A V_i(r_i, \varphi_i) - \mathbf{I} \omega' \right\} \phi'^{CM}, \quad (3.22a)$$

$$\mathbf{I}_0 \phi'^{CM} = 0, \quad (3.22b)$$

where $\omega' = R^{-1} \omega R$.

If ω' is treated classically, the moment of inertia is obtained from the calculation of the second order perturbation with respect to ω' . It should be noted that the unperturbed wave function corresponds to the $\phi_i^M(x, 0)$ in (3.15).

Thus, it has been made clear that the model wave function must be introduced in the representation (3.4). In this representation the relation among the various methods with the deformed potential model becomes evident.

§ 4. Expression for the moment of inertia

Now, let us consider the expression for the moment of inertia \mathcal{J} in the representation (3.4). The following operator equations can be set up in this representation,

$$\frac{i}{\hbar} \left[H_0 - \frac{1}{2\mathcal{J}} \mathbf{I}_0^2, \vartheta \right] = \frac{1}{\mathcal{J}} \mathbf{I}, \quad (4.1)$$

$$[\mathbf{I}, \vartheta] = -i\hbar, \quad (4.2)$$

where ϑ is assumed to have the same functional form as that defined through (2.10) and (2.11).

Let us solve these operator equations by using $\phi_i(x, t)$ and $\phi_i^M(x, t)$. Then we obtain the usual expressions for \mathcal{J} respectively^{5)10)*}

$$\mathcal{J} = 2 \sum_{j \neq i} \frac{|\langle \phi_j | \mathbf{I} | \phi_i \rangle|^2}{\epsilon_j - \epsilon_i} \quad (4.3a)$$

$$\mathcal{J}^M = 2 \sum_{j \neq i} \frac{|\langle \phi_j^M | \mathbf{I} | \phi_i^M \rangle|^2}{\epsilon_j^M - \epsilon_i^M}. \quad (4.3b)$$

The model wave function ϕ_i , which is not an eigenstate of the total angular momentum, seems at a glance to give the meaning to (4.3a). However, (4.3a) has no meaning actually, because

$$\langle \phi_j | \mathbf{I} | \phi_i \rangle = 0 \quad \text{for } i \neq j$$

as easily seen from (3.10).

On the other hand, the model wave function ϕ_i^M in the deformed potential model has non-diagonal matrix elements of \mathbf{I} , so that (4.3b) has a meaning. This circumstance results from the existence of the deviation terms in (3.16), which are due to the fact that ϕ_i^M is not exactly an eigenstate of $(H_0 - (1/2)\gamma)\mathbf{I}_0^2$. Therefore, the relation between the real moment of inertia and γ^M in (4.3b) becomes obscure. From this point of view, the validity of the cranking model formula for \mathcal{J} may be doubtful.

If we obtain the expression for γ being meaningful even for ϕ_i , it may give the real moment of inertia in principle.

Another interesting expression for γ , which has significance even for ϕ_i , is derived as follows. Operating \mathbf{I}^2 on the left through (3.11a), we have

$$\mathbf{I}^2 \left(H_0 - \frac{1}{2} \gamma \mathbf{I}_0^2 \right) \phi_i = \epsilon_i \mathbf{I}^2 \phi_i. \quad (4.4)$$

Therefore we get

$$\langle i | \mathbf{I}^2 H_0 | i \rangle - \frac{1}{2} \gamma \langle i | \mathbf{I}^2 \mathbf{I}_0^2 | i \rangle = \epsilon_i \langle i | \mathbf{I}^2 | i \rangle.$$

Taking into account the relation

This expression (4.3) is derived through the following relation,

$$\langle \phi_i | [H_0 - (1/2)\gamma]\mathbf{I}^2 | \phi_i \rangle = 2 \sum_{j \neq i} \frac{\langle \phi_i | H_0 - (1/2)\gamma | \phi_j \rangle \langle \phi_j | \mathbf{I}^2 | \phi_i \rangle}{\epsilon_j - \epsilon_i}.$$

Because explicit use is made of matrix multiplication, it must be guaranteed actually that the state $\nu\phi_k$ remains in the space of the set of eigenfunctions ϕ_k . Therefore, strictly speaking, ν in the above relation must be replaced by the periodic function in ν with period 2π , e.g., $\sin \nu$, and the expression of \mathcal{J} , (4.3), must be modified slightly. However, as discussed by Lipkin et al.¹¹⁾ the expression (4.3) holds approximately in the case where the deformed potential model wave function is used.

$$\epsilon_i = \langle i | H_0 | i \rangle - \frac{1}{2\mathcal{J}} \langle i | \mathbf{L}_0^2 | i \rangle,$$

we obtain

$$\mathcal{J} = \frac{1}{2} \frac{\langle i | \mathbf{L}_0^2 | i \rangle - \langle i | \mathbf{L}_0^2 | i \rangle \langle i | \mathbf{L}^2 | i \rangle}{\langle i | \mathbf{L}^2 H_0 | i \rangle - \langle i | \mathbf{L}^2 | i \rangle \langle i | H_0 | i \rangle}. \quad (4.5)$$

Since ϕ_i satisfies the subsidiary condition (3.11b), (4.5) becomes

$$\mathcal{J} = \frac{1}{2} \frac{\langle i | \mathbf{L}^4 | i \rangle - \langle i | \mathbf{L}^2 | i \rangle^2}{\langle i | \mathbf{L}^2 H_0 | i \rangle - \langle i | \mathbf{L}^2 | i \rangle \langle i | H_0 | i \rangle}. \quad (4.6)$$

This is just Skyrme's formula, which reduces to that obtained by Peierls and Yoccoz under an appropriate approximation. It is to be remarked that if we use the real wave function ϕ_{iB} , this expression has no meaning.

Now, it is clear that the expression (4.6) is more reliable than the cranking model formula (4.3).

§ 5. Concluding remarks

It has become evident that the introduction of the model wave function is necessary so long as we calculate the moment of inertia by the equation (4.3) or (4.6). The introduction of this model wave function is possible only in the representation (3.4). The relation among the various methods with the deformed potential model does not become clear till the consideration is taken into account in this representation.

It should be stressed here that the above discussions were possible only on the conditions (2.10) and (2.11) which assure the separation of the real Hamiltonian into rotational and intrinsic parts. In other words, it is just this condition that guarantees the validity of the methods with the deformed potential model. Furthermore, it is to be remarked that the model wave function should be restricted in the form of ϕ_i in (3.10) and the expression having significance even for ϕ_i should be used in calculating the moment of inertia.

As mentioned in § 1, the conditions for the separation of nuclear Hamiltonian are connected essentially with the interactions between nucleons. Thus the significance and the importance of these conditions will be made clear in the next stage of the study of nuclear structure, in the sense that they may show a formal condition which is to be satisfied by the actual nuclear interactions. It will be the most profound problem to make clear this circumstance concerning the investigation of the origins of nuclear deformation and other fundamental properties of nuclei.

References

- 1) F. Villars, *Annual Review of Nuclear Science* **7** (1957), 185.
Earlier references are given in this paper.
- 2) R. E. Peierls and J. Yoccoz, *Proc. Phys. Soc.* **A70** (1957), 381.
J. Yoccoz, *Proc. Phys. Soc.* **A70** (1957), 388.
- 3) J. J. Griffin and J. A. Wheeler, *Phys. Rev.* **108** (1957), 311.
- 4) D. R. Inglis, *Phys. Rev.* **96** (1954), 1059; *Phys. Rev.* **103** (1956), 1786.
- 5) H. J. Lipkin, A. de-Shalit and I. Talmi, *Phys. Rev.* **103** (1956), 1773.
- 6) S. A. Moszkowski, *Phys. Rev.* **103** (1956), 1328.
- 7) A. Bohr and B. R. Mottelson, *Dan. Mat. Fys. Medd.* **30** (1955), No. 1.
- 8) T. H. R. Skyrme, *Proceeding. Roy. Soc.* **239** (1957), 399; *Proc. Phys. Soc.* **A70** (1957), 433.
- 9) H. J. Lipkin, *Proceeding Rehovoth Conference on Nuclear Structure*, 144,
(North Holland Publishing Company, Amsterdam, 1958)
H. J. Lipkin and S. Goldstein, *Nuclear Physics* **7** (1958), 184.
- 10) A. Bohr and B. R. Mottelson, preprint.
- 11) R. Nataf, *Nuclear Phys.* **2** (1956/57), 497.
- 12) H. J. Lipkin, A. de-Shalit and I. Talmi, *Nuovo Cimento* **2** (1955), 773.
- 13) T. Tamura, *Nuovo Cimento* **4** (1956), 713.
- 14) S. Nagata, R. Tamagaki, S. Amai and T. Marumori, *Prog. Theor. Phys.* **19** (1958), 495.
- 15) S. Hayakawa and T. Marumori, *Prog. Theor. Phys.* **18** (1957), 396.

Letters to the Editor

The opinions expressed in these columns do not necessarily reflect those of the Board of Editors. Communications should be submitted in duplicate and should be held to within 100 lines (pica type) on standard size letter paper (approx. 21×30 cm.), so that each letter will be arranged into two pages when printed. Do not forget to count in enough space for formulas, figures or tables.

Influence of the Exchange and Correlation on the Electron-Phonon Interaction in Metals

Tomiyuki Toya

*The Research Institute for Catalysis,
Hokkaido University, Sapporo*

August 21, 1958

Recently Bardeen and Pines ($B-P$)¹⁾ investigated the role of electron-electron interactions in the electron-phonon interaction in metals by extending the Bohm-Pines collective description method. The conclusion is that the exchange and correlation effects on the dispersion relation $\omega(q)$ and on the effective matrix element $V_p(q)$ for electron-phonon interaction is negligible for small wave number vector q of a phonon.

The conclusion might, however, be questioned as follows. According to $B-P$ the effective matrix element $V_p(q)$ is given without the exchange and correlation effects as

$$V_p(q) = V_i(q) \{1 + 8\pi e^2/q^2 \times \sum (k) (E_0(k+q) - E_0(k))^{-1}\}^{-1}, \quad (1)$$

where $V_i(q)$ is the matrix element of the electron-phonon interaction, and $E_0(k)$ is the energy of an electron of the wave number vector k , including the coulomb energy due to other electrons. The com-

pressibility $\kappa = \{c_{11} - 2/3(c_{11} - c_{12})\}^{-1}$ is now formulated²⁾ in terms of the above $V_p(q)$ by expressing elastic constants c_{11} and $c_{11} - c_{12}$ by the dispersion relation $\omega(q)$, which includes $V_p(q)$, in the long wavelength limits; the result is

$$(1/\kappa)_H = N \{ -2(V(r_s) - E_0(k=0)) + 2/3 \cdot \zeta_0 + 0.2 e^2/r_s \}, \quad (2)$$

where N is the number of valence electrons in unit volume, ζ_0 the Fermi energy, $V(r_s)$ the Hartree field at the surface of the atomic sphere and r_s its radius.

The compressibility is, on the other hand, derived from the cohesive energy E_c according to the Wigner-Seitz theory, as³⁾

$$\begin{aligned} 1/\kappa &= (1/12\pi r_s) d^2 E_c / dr_s^2 \\ &= (1/\kappa)_H - N [0.204 e^2/r_s \\ &\quad + 0.064 e^2/r_s \{1 + (1 + 5.1 a_0/r_s)^{-1}\} \\ &\quad \times (1 + 5.1 a_0/r_s)^{-2}], \end{aligned} \quad (3)$$

referring to the equilibrium condition $dE_c/dr_s = 0$, and the relation, $dE_0/dr_s = 3/r_s \{ (V(r_s) - E_0(k=0)) - 0.2 e^2/r_s \}$, derived by Froelich³⁾ and Bardeen.⁴⁾ In eq. (3), a_0 is the Bohr radius and the first term $(1/\kappa)_H$ is given by eq. (2), while the second term, arising from the exchange and correlation terms in E_c , is not at all negligible; e. g. for Na the first or second terms is $1.39 \cdot 10^{11}$ dynes/cm² or $0.67 \cdot 10^{11}$ dynes/cm², respectively and hence $1/\kappa = 0.72 \cdot 10^{11}$ dynes/cm² or $0.80 \cdot 10^{11}$ dynes/cm², taking into account the contri-

bution from the ion-core exchange repulsion, compared with the experimental value $\sim 0.80 \cdot 10^{11}$ dynes/cm².

We are thus led to the conclusion that $\omega(q)$ and $V_p(q)$ must include the exchange and correlation terms even for small q at variance with that of $B-P$.

Present author has included the exchange and correlation effect in the electron-phonon interaction by extending the Hartree-Fock method and taking the configurational interaction into account by Maecke's method; the effective matrix element is thus found to be,

$$V_p(q) = V_i(q) \{ 1 + (8\pi e^2/q^2 - 2B(q)\pi e^2/k_F^2) \times \sum (E(k+q) - E(k))^{-1} \}^{-1}, \quad (4)$$

instead of eq. (1), where $E(k)$ is the energy of an electron including the exchange and correlation energy due to other electrons and k_F the wave number of an electron at the Fermi surface. The term $B(q)$, which distinguishes the above equation from eq. (1), is rather complicated but is estimated approximately to be $1.25 - 0.25 q^2/q_{\max}^2$ for Na.

$B(q)$ is physically interpreted as giving the exchange and correlation effects which reduces the shielding field as follows. The charge density of electrons is redistributed by crystal deformation to shield the perturbing ionic field. Electrons tend, however, to be apart from each other by exchange and correlation interactions to reduce the shielding effect compared with the case of their absence. The shielding effect is further reduced through the term $\sum (E(k+q) - E(k))^{-1}$ in eq. (4), which replaces $\sum (E_0(k+q) - E_0(k))^{-1}$ in eq. (1), as seen from the approximate relation,

$$\sum (E(k+q) - E(k))^{-1} = D/D_0 \sum (E_0(k+q) - E_0(k))^{-1}, \quad (5)$$

where D or D_0 is the density of states with or without the exchange and correlation energies and D/D_0 for monovalent metals.

The contribution of the exchange and correlation energies to the compressibility is derived from the dispersion relation $\omega(q)$ and eq. (4) as

$$-N \{ B(q=0) \cdot 0.204 e^2/r_s + 2 \cdot 3 \cdot \frac{e^2}{\epsilon_0} (D/D_0 - 1) \}, \quad (6)$$

which is numerically equal to the second term of eq. (3) for Na with $D/D_0 = 0.90^{10}$ and $B(q=0) = 1.25$.

The validity of eq. (4) for $q \approx q_{\max}$ is on the other hand examined as follows. For the mode $q = (\pi/r_0, 0, 0)$, $e_q = (1, 0, 0)$ or $q = (0, \pi/r_0, 0)$, $e_q = (1, 0, 0)$ in the metals of b. c. c. structure ($2r_0$: lattice const.), each ion is displaced alternately by $+\partial R$ or $-\partial R$ in the direction of x -axis from the equilibrium position, and hence the contributions from the valence electrons to the adiabatic potential may be calculated by the cell method,¹¹ admitting that the valence-electron coulomb energy is approximately balanced by the exchange and correlation energies just as in perfect crystals. The circular frequency is thus found to be $1.64 (Ne^2/M)^{1/2}$ (M : the atomic mass), in comparison with $1.60 (Ne^2/M)^{1/2}$ as obtained by the dispersion relation and eq. (4).¹² The appropriate adiabatic potential amount to 50% larger, if exchange and correlation effects are ignored.

- 1) J. Bardeen and D. Pines, Phys. Rev. **99** (1955), 1140.
- 2) T. Taya, Busseiron Kenkyu **60** (1953), 129.
- 3) H. Froelich, Proc. Roy. Soc. **158** (1937), 97.
- 4) J. Bardeen, J. Chem. Phys. **6** (1938), 372.
- 5) D. Pines, Phys. Rev. **92** (1953), 39.

Electrical Conductivity of Monovalent Metals

Tomiyuki Toya

*The Research Institute for Catalysis,
Hokkaido University, Sapporo*

August 21, 1958

The interaction constants C of monovalent metals are evaluated from the effective matrix element of electron-phonon interaction derived by self-consistent field method, taking the exchange and correlation effects into account.

With the wave function,

$$\varphi(\mathbf{k}) = \exp(i\mathbf{k}\mathbf{r}) u_0(\mathbf{r}), \quad (1)$$

of an electron of wave number vector \mathbf{k} and of its Hartree energy $E_0(\mathbf{k}) = E_0 + \hbar^2/2m \cdot k^2$, the effective matrix element $V_p^0(\mathbf{q})$ of electron-phonon interaction is given as¹⁾

$$V_p^0(\mathbf{q}) = V_i^0(\mathbf{q}) (1 + S(\mathbf{q}))^{-1}, \quad (2)$$

where

$$S(\mathbf{q}) = (8\pi e^2/q^2 - 2B(\mathbf{q})\pi e^2/k_F) \times D/D_0 \sum_{\mathbf{k}} (E_0(\mathbf{k} + \mathbf{q}) - E_0(\mathbf{k}))^{-1}, \quad (3)$$

and,²⁾

$$V_i^0(\mathbf{q}) = iN^{-1/2}(\mathbf{e}_q, \mathbf{q}) \{4\pi Ne^2/q^2 + \gamma(V(r_s) - E_0(\mathbf{k}=0))\} g(\mathbf{x}); \quad (4)$$

\mathbf{e}_q is the displacement vector of ions, $g(\mathbf{x}) = 3(\sin \mathbf{x} - \mathbf{x} \cos \mathbf{x})/\mathbf{x}^3$, $\mathbf{x} = r_s \mathbf{q}$, $V(r_s)$ is the Hartree potential energy at the surface of atomic sphere of radius r_s and $\gamma = u_0^2(r_s)$. D or D_0 is the density of states at the Fermi surface with or without exchange and correlation energies, and the term $B(\mathbf{q})$ in eq. (3) allows for the effect

of the exchange and correlation energies on the screening field, as discussed in the foregoing letter.¹⁾

Wave function (1), on which eqs. (2), (3) and (4) are based, is, however, known as a poor approximation in the case of Li; the approximation might be improved by using the function,

$$\varphi(\mathbf{k}) = \exp(i\mathbf{k}\mathbf{r}) \{u_0(\mathbf{r}) + i(\mathbf{k}\mathbf{r})v(\mathbf{r})\}, \quad (5)$$

and its energy $E_0(\mathbf{k}) = E_0 + \alpha \hbar^2/2m \cdot k^2$, for calculating the matrix elements, where α is given as

$$\alpha - \gamma = u_0(r_s) (dv/dr)_{r=r_s} r_s.$$

We have thus, approximately,

$$V_i(\mathbf{q}) = V_i^0(\mathbf{q}) + (k/k_F)^2 \sin^2(\mathbf{q}\mathbf{k}) V_i'(\mathbf{q}), \quad (6)$$

where $V_i^0(\mathbf{q})$ is given by eq. (4) and

$$V_i'(\mathbf{q}) = iN^{-1/2}(\gamma - \alpha)(\mathbf{e}_q, \mathbf{q}) (\hbar^2/2m \cdot k_F^2) \times 6/x^2 (g(\mathbf{x}) - \sin \mathbf{x}/\mathbf{x}),$$

and k_F is the wave number at the Fermi surface. The effective matrix element is derived self-consistently as,

$$V_p(\mathbf{q}) = (V_i^0(\mathbf{q}) - V_i'(\mathbf{q})T(\mathbf{q}))(1 + S(\mathbf{q}))^{-1} + (k/k_F)^2 \sin^2(\mathbf{q}\mathbf{k}) V_i'(\mathbf{q}), \quad (7)$$

where

$$T(\mathbf{q}) = (8\pi e^2/q^2 - 2B\pi e^2/k_F^2) \times D/D_0 \sum_{\mathbf{k}} (k/k_F)^2 (E_0(\mathbf{k} + \mathbf{q}) - E_0(\mathbf{q}))^{-1}.$$

The terms omitted in eqs. (6) and (7) are proportional to q^3 and are negligible for $q \ll q_{\max}$.

The interaction constants C are now evaluated as follows: $B(\mathbf{q})$ are determined, when we assume each of them, respectively,

Table I

	Li	Na	K	Rb	Cs	Cu
$\sigma \times 10^{-16}$ (in c. g. s. unit) ($T^\circ K$)	10.65 (250)	22.9 (150)	18.2 (80)	8.6 (50)	6.0 (35)	53.8 (273)
r_s (in atomic unit)	3.22	3.96	4.87	5.18	5.57	2.68
$\alpha = m/m^*$	0.69	1.0	1.0	1.0	1.0	1.0
$\gamma - \alpha^{(1)}$	0.31	0	0	0	0	0
ζ_0 (eV)	3.32	3.17	2.10	1.86	1.61	6.95
$B(q=0)$	1.32	1.25	1.21	1.18	1.18	1.37
$D/D_0^{(b)}$	0.837	0.90	0.935	0.945	0.965	0.874
$V-E_0$ (eV) ^(c)	-0.80	0.08	-0.02	0.03	0.00	3.7
C/ζ_0 (from eq. (2) or eq. (7))	~ 1.3	1.08	1.12	1.20	1.20	1.51
C/ζ_0 (Hartee approx.) ^(d)	0.72	0.75	0.77	0.80	0.82	1.12
$\theta^\circ K$	330~360	180~202	114	65	45	320
C/ζ_0 (experiment)	1.45~1.6	0.98~1.08	0.98	1.32	1.29	1.51

a) We assumed $\alpha = \gamma = 1$ for metals except Li. b) D. Pines, *Phys. Rev.* 92 (1954), 626. c) F. S. Ham, *Solid State Physics*, Vol. I (1955), 185. For the value of Cu, see the following letter in this issue. d) See Ref. (2).

being approximately constant at $q=0$, and equating the second term of eqs. (4) and (6) of the foregoing letter, which both are the contribution of the exchange and correlation energies to the compressibility.¹⁾ The *Umklapp* processes are taken into account in the same way as done by Bardeen.²⁾ Table I shows the calculated values of C and associated quantities in comparison with its experimental values C_{exp} , which have been worked out from the observed conductivities $\sigma(T)$ according to the equation,³⁾

$$C_{\text{exp}}^2 = \text{const.} \cdot k_F (dE/dk)_F^2 M \theta (T\sigma)^{-1},$$

where M is the atomic mass, θ the Debye temperature, and $(dE/dk)_F = D/D_0 (dE_0/dk)_F$.

As seen from Table I, the theoretical values are considerably improved compared with those of the Hartree approximation,²⁾ by taking the exchange and correlation effects into account.

- 1) T. Tova, *Bulleten Kerkva* 59 (1953), 179; *ibid.*, 99 (1955), 33; *Prog. Theor. Phys.* 19 (1958), this issue.
- 2) J. Bardeen, *Phys. Rev.* 52 (1937), 688.
- 3) A. Sommerfeld & H. A. Bethe, *Handbuch der Physik*, Vol. 24, II (1933), 513. The conclusion of E. J. Nist (*Phys. Rev.* 99 (1935), 1735) or R. Barrie (*ibid.*, 103 (1956), 1581) is not correct. Cf. M. Tsuji, *J. Phys. Soc. Japan*, 12 (1957), 828.

Normal Vibrations of Copper

Tomiyuki Toya

The Research Institute for Catalysis,
Hokkaido University, Sapporo

August 21, 1958

Frequency vs wave vector relations for Cu are calculated in the (100), (110)

and (111) direction on the basis of the dispersion relation,

$$M\omega(q)^2 = - \sum_{k,h} \{ V_p^*(q+K_h) \times V_i(q+K_h) (E(k+q+K_h) - E(k)) + V_i^*(q+K_h) \times V_p(q+K_h) / (E(k-q-K_h) - E(k)) \} + 4\pi Ne^2/3 + H_c + H_r, \quad (1)$$

which have previously been derived from the first principles by the present author for the calculation of the normal vibrations of Na.¹⁾ In eq. (1), M is the atomic mass, ω circular frequency of the longitudinal or transversal lattice vibration of the wave number vector q in the principal directions, K_h a vector of the reciprocal lattice space, $V_i(q+K_h)$ and $V_p(q+K_h)$ matrix elements of electron-phonon and effective electron-phonon interactions respectively, $E(k)$ the energy of an electron of wave number k , and N the number of valence electrons in unit volume. The third and fourth terms, H_c and H_r , are the contributions of potential energies of the coulomb and exchange repulsions between ion-cores respectively.

The first term represents the contribution from the electron-phonon interactions, including electron-electron interactions. The summation extends over the valence electrons, and *Umklapp* processes $K_h \neq 0$ as well. The second term covers the second order perturbation of the ionic field by lattice deformation.

The first term in eq. (1) is evaluated assuming $\alpha = m/m^* = 1$, $\gamma = u_0^2(r_s) = 1$ and $V(r_s) - E(k=0) = 3.7$ eV. (notations are the same as those in the foregoing letters²⁾). These assumptions are reasonable except for electrons near the Fermi surface on the

ground of the recent theoretical calculations of the electronic structure of Cu. We have, besides, $D/D_0 = 0.875$ and $B = 1.37$, similarly as in the foregoing letter.²⁾ Correction term $2.90 \cdot 10^{-4} r_s^4 (q_x^4 + q_y^4 + q_z^4) - 2.06 \cdot 10^{-10} r_s^4 |q|^4$ is added to $g(x)$ in the expressions V_i and V_p , as to satisfy the condition $V_i(K_h) = V_p(K_h) = 0$ for $K_h = \pi/r_0$ (1, 1, 1) and $K_h = \pi/r_0$ (2, 0, 0) ($2r_0$; the lattice const.). H_c is readily computed by Ewald's method,⁴⁾ its contribution to the elastic constants $c_{11} - c_{12}$ and c_{44} being found identical with the results of Fuchs,⁵⁾ as determined in the long wavelength limits of transversal vibrations. We compute H_r assuming the exchange repulsions $U(r)$ between ion-cores to be the central force potentials of the Born-Mayer type, as

$$U(r) = 0.120 \exp 11.5 (r_e - r)/r_e \text{ eV}, \quad (2)$$

where $r_e = \sqrt{2} r_0$ is the distance between nearest neighbours in the equilibrium configuration of the lattice.

Frequency *vs* wave vector relations are evaluated with the above values, as shown by solid curves in Figs. (1), (2) and (3).

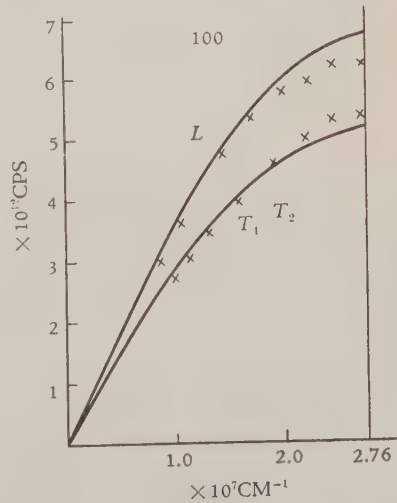


Fig. 1

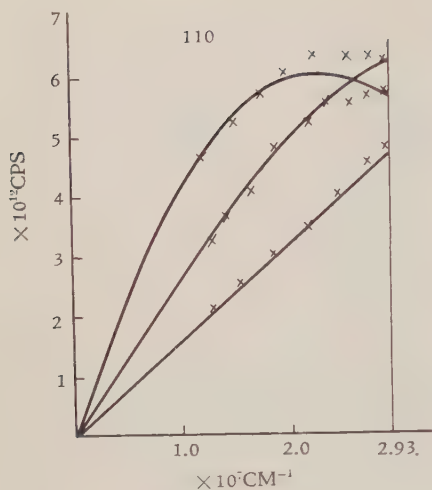


Fig. 2

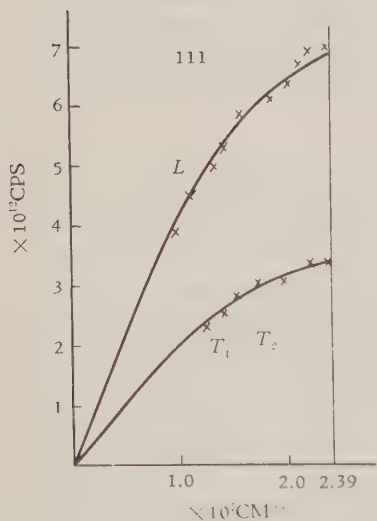


Fig. 3

Crosses in the figures show the experimental results of Jacobsen⁽⁶⁾ determined from the temperature diffuse scattering of X-rays; two constants of eq. (2) have been adjusted to the experimental values at the largest q in Fig. (3).

The elastic constants c_{11} , c_{44} and $c_{11}-c_{12}$ are estimated approximately to be 14.0, 8.2 and 4.6 in units of 10^{11} dynes/cm² respectively. The electronic term compensate 80% or 20% of the contribution from

the coulomb terms (the second and third terms of eq. (1)) to c_{44} or $c_{11}-c_{12}$.⁽⁷⁾ The value of c_{11} is somewhat small compared with the experimental value. The discrepancy might be attributed to the simple assumptions for the electrons near the Fermi surface as well as for ion-core exchange repulsions given by eq. (2).

The author expresses his sincere thanks to Professor J. Horiuti for his continual encouragement and interest shown to the present three works.

- 1) T. Toya, *Busseiron Kenkyu* **60** (1953), 129; cf. also J. Bardeen and D. Pines, *Phys. Rev.* **99** (1955), 1140 and T. Toya, *J. Research Institute for Catalysis Vol. 6* (to be published).
- 2) T. Toya, *Prog. Theor. Phys.* **19** (1958), in this issue.
- 3) K. Kambe, *Phys. Rev.* **99** (1955), 419; M. Fukuchi, *Prog. Theor. Phys.* **16** (1956), 222; P. J. Howrth, *Proc. Roy. Soc. A* **220** (1953), 513; *Phys. Rev.* **99** (1955), 46.
- 4) P. P. Ewald, *Nachr. Akad. Wiss. Göttingen, N. E. II?* **3** (1938), 55. E. W. Kellerman, *Phil. Trans. Roy. Soc. London A* **238** (1940), 513.
- 5) K. Fuchs, *Proc. Roy. Soc. A* **151** (1935), 585; **A 153** (1935), 622; **A 157** (1936), 444.
- 6) E. H. Jacobsen, *Phys. Rev.* **97** (1955), 654.
- 7) Cf. C. Zener, *Acta cryst.* **3** (1950), 346; H. B. Huntington, *Phys. Rev.* **91** (1953), 1092. In the case of Na, the contribution from the first term to shear moduli is negligible as assumed by Fuchs (cf. ref. (1)).

A Note on the Elastic Scattering of High Energy Deuterons by Complex Nuclei

Yoshihiko Nishida and Hiromi Tanaka

Department of Physics, Saga University, Saga

September, 18, 1958

Recently it was pointed out by Sakamoto and Sasakawa¹⁾ that a faithful application

of the impulse approximation to elastic scattering of high energy deuterons by nucleons leads to a more excellent agreement with experiment than Chew's original result.²⁾ As for elastic scattering by complex nuclei, however, the impulse approximation will not be so good. The analysis by Baldwin et al.³⁾ shows that as smaller scattering angles the differential cross section by the impulse approximation is fairly larger than the observed one, but with increasing scattering angles the former becomes conversely much smaller than the latter. The success of the impulse approximation in nucleon-deuteron elastic scattering suggests that most part of this large disparity between theory and experiment may be due to the simultaneous scattering of both nucleons of deuterons by nuclei. Now, Akhieser and Sitenko⁴⁾ (hereafter referred to as $A-S$) recently proposed the theory of diffractive scattering of high energy deuterons by nuclei. Their theory corresponds to a generalization of the optical model to the deuteron-scattering. It is very clear-cut in principle and the effect of simultaneous scattering is straightforwardly introduced. Therefore, we may consider their theory as an improvement on the impulse approximation. Then, it comes into question to investigate how far theory is available in practical analyses. In this note we shall report some results obtained by applying the $A-S$ theory to the elastic scattering of 156 MeV deuterons by carbon.⁵⁾ On the other hand, the effect of simultaneous scattering was independently discussed also by Stapp⁶⁾ with a quite different idea. Therefore, the following discussion will naturally be developed in close contact with his theory.

First of all, it might be pointed out

that although Stapp's theory is developed in a laborious manner compared with the $A-S$ theory, both theories lead to nearly the same results. The W. K. B. approximation on which the $A-S$ theory is based, well reproduces the nucleon-nucleus scattering data at high energies. So, it will be permitted within the frame of the $A-S$ theory to replace the nucleon-nucleus scattering amplitude by a suitable empirical formula. To compare with the result of Stapp's theory, it is desirable to use the same empirical formula for the nucleon-nucleus scattering amplitude as he did. Indeed, we also use the form obtained by Greider³⁾ to fit the cross section data of Strauch and Titus for 96 MeV protons on carbon. The scattering amplitude for simultaneous scattering is calculated with the assumption that the phase in the nucleon-nucleus scattering amplitude is angle-independent and with the Gaussian type of S -state deuteron wave function. When the spin-dependent effect is omitted, the result shows that $I(x)$ and $I(x, y)$ appearing in the scattering amplitude in Stapp's theory have only to be replaced by $I(x)$ and $I(x, y)$, respectively. There are simple relations between them:

$$I(x) \approx \Gamma(x) [R_d^2 + 4x \sin^2 \frac{1}{2} \theta / R_d^2]^{1/2},$$

$$I(x, y) \approx \Gamma(x, y)$$

$$\times [R_d^2 + 2(x+y) \sin^2 \frac{1}{2} \theta / R_d^2]^{1/2}.$$

Here $R_d = 2.64$ fermi is the deuteron radius and θ is the deuteron-scattering angle. However, the effect of this substitution is easily seen to be negligibly small at the scattering angles less than 30° . Therefore, we shall find it convenient to use the $A-S$ theory for a practical calculation. The differential cross section

obtained using the $A-S$ theory is compared with the result by using the impulse approximation in Fig. 1.

Next, we calculate the differential cross section using a more familiar deuteron wave function $\varphi(r) = \sqrt{\alpha/2\pi} e^{-\alpha r}/r$ to make a detailed comparison with experimental data. Difficulty in the analytical integration in this case be avoided by making a suitable approximation in the integrand:

$$4\alpha/k \arctan k/4\alpha \approx p_1 \exp(-\lambda k^2/2) + p_2 \exp(-\mu k^2/2) + p_3 \exp(-\nu k^2/2)$$



- • • • • Experimental points.
- Impuls approximation for Gaussian deuteron wave function.
- Theoretical curve for the Gaussian one.
- · - · - Impuls approximation for the familiar one.
- Theoretical curve for the familiar one.

Fig. 1. Angular distribution of 156 MeV

$$\begin{aligned} \text{where } p_1 &= 0.59, p_2 = 0.10, p_3 = 0.31 \\ \lambda &= 0.8 \text{ fermi}^2, \mu = 0.2 \text{ fermi}^2, \\ \nu &= 0.01 \text{ fermi}^2. \end{aligned}$$

The theoretical differential cross section is in good agreement with the observed one at smaller angles, and even at larger angles the former is smaller than the latter, at worst by the factor of 1.4 only. About the detail, see also Fig. 1. It might be mentioned that the angular distribution at smaller angles is insensitive to the form of sticking factor.

Finally, the comments will be made on the assumption of angle-independent phase and the limit of applicability of the $A-S$ theory. In the W. K. B. approximation, the phase of the nucleon-nucleus scattering amplitude in this case increases very slowly with the scattering angle for the angles less than about 30° . So, the above assumption will be a good approximation in this region. This means that only for the deuteron-scattering angles less than about 15° the assumption of constant phase is a reasonable approximation. It will be understandable that for the deuteron-scattering angles over about 15° , the theoretical differential cross section deviates from the observed one. In this region, however, the W. K. B. approximation for the nucleon-nucleus scattering will be less valid. So, it will be unsuitable for improvement of the situation to use the angle-dependent phase by the W. K. B. approximation. We would rather consider that this region lies outside the limit of applicability of the $A-S$ theory. To obtain a satisfactory agreement with experiment including the polarization effect in this region, it is necessary to apply an improved approximation on the W. K. B. one. It might be

emphasized that the $A-S$ theory (and also the Stapp theory) will be available enough to improve the impulse approximation only for smaller scattering angles.

- 1) Y. Sakamoto and T. Sasakawa, *Prog. Theor. Phys.* **19** (1958), 745.
- 2) G. F. Chew, *Phys.* **74** (1948), 809; **80** (1950), 196.
- 3) Baldwin, Chamberlain, Segrè, Tripp, Wiegand and Ypsilantis, *Phys. Rev.* **103** (1956), 1502.
- 4) A. I. Akhiezer and A. G. Sitenko, *Phys. Rev.* **106** (1957), 1236.
- 5) H. P. Stapp, *Phys. Rev.* **107** (1957), 607.
- 6) K. Greider, unpublished result. See (60) (61) in Stapp's paper.

Some Remarks on the Scattering Problem

Masakuni Ida

*Department of Physics, University of Kyoto,
Kyoto*

October 2, 1958

Invalidity of Levinson's theorem in meson scattering.

We know Levinson's theorem¹⁾ in quantum mechanics which states under certain conditions,

$$\delta(\infty) - \delta(0) = -m\pi,$$

where $\delta(p)$ is the scattering phase shift for momentum p of the scattering particle and m the number of bound states. It is often applied to the problem of meson scattering. But it is essential in the derivation of this theorem that the scattering is due to a local, velocity independent

potential. Since the scattering of mesons cannot be described as the one due to such a potential, it is doubtful whether the theorem keeps its validity for meson scattering. Indeed, we can show that this is not the case.

In a simple model presented by Dyson,²⁾ assuming that the square of the cut-off factor is of the order of $1/p^{1+s}$, with $s > 0$, as $p \rightarrow \infty$, we obtain for $N-\theta$ scattering the relation

$$\tau \equiv \frac{1}{\pi} \{ \delta(\infty) - \delta(0) \} = n + 1 - m,^*$$

when n is the number of zeros of the functions $h(z)$ defined below:

$$h(z) = \sum_{i=1}^m \frac{\lambda_{B_i}}{\omega_{B_i} - z} + \frac{1}{\pi} \int_{\mu}^{\infty} \frac{d\omega_q}{qv^2(q)} \frac{\sin^2 \delta(q)}{\omega_q - z},$$

$$\lambda_{B_i} > 0, \quad \omega_{B_i} < \mu, \quad \omega = (\mu^2 + q^2)^{1/2}.$$

λ_{B_i} defines the effective coupling of the i -th bound state B_i with the energy $m_N + \omega_{B_i}$ to the $N-\theta$ system. m_N and μ are the masses of the N and θ particles, respectively. The cut-off factor, $v(q)$, is normalized in such a way that $v(0) = 1$.

The parameter $\sigma = m + \tau$, indicating the deviation from Levinson's theorem, determines the number of zeros of $h(z)$, accordingly the form of the solutions of the Low equation.³⁾ In other words, σ is the parameter concerned with the selection of redundant solutions⁴⁾ of the Low equation for this model. σ equals the number, $n+1$, of the bare V particles which cannot be directly observed. We notice that σ itself is an observable quantity.

* When $h(\mu) = \infty$, then $\tau = n + 1/2 - m$. A resonance occurs at zero kinetic energy. For the sake of simplicity we do not consider this case in this letter.

Levinson's theorem is in general invalid for the charged scalar meson theory⁽¹⁾ also, and perhaps for the symmetric pseudo-scalar one.⁽²⁾ The theorem loses its kinematical character in the case of meson scattering.

Positive and negative resonances.

We classify the resonances into two types, positive and negative, according to the sign of $d\delta(p)/dp$ at the resonant energies. If we denote the number of resonances of respective types by r_p and r_u , then the relation

$$\tau = r_p - r_u$$

holds generally under the elastic approximation and the assumption that the cut-off factor damps sufficiently for high energies.

An example of negative resonance is obtained for the charged scalar meson theory with the one-meson approximation. In the Serber-Lee solution⁽¹⁾ there is a bound state in the $\pi^+ + p$ system for the square of the meson-nucleon coupling constant $\lambda_N > 1$. Corresponding to this bound state there appears a resonance of the negative type at the meson energy $\omega = \lambda_N/\mu$.

Resonances of the two types have some mutually contrasting properties:

a) When we analytically continue the S -matrix downwards on the right side of the branch point at $\omega = \mu$, there usually appears in the lower half of the complex ω plane a pole of $S(\omega)$ corresponding to a positive resonance,⁽³⁾ but it seems rather usual that there is no such pole for a negative one.

b) In the resonant scattering of a wave packet, the packet is delayed by a positive resonance, but is advanced by a negative one.

c) The width of a negative resonance has a lower bound in order to be consistent with the requirement of causality.

The well-known 3-3 resonance in the low energy $\pi - N$ collision is of the positive type. It may be interesting whether negative resonances appear in the higher energy region.

Details of this letter will be reported before long in this journal.

- 1) N. Levinson, Kgl. Danske Videnskab. Selskab, Math.-fys. Medd. **25** No. 9 (1949).
R. Jost, Helv. Phys. Acta **20** (1947), 256.
- 2) F. J. Dyson, Phys. Rev. **106** (1957), 157.
- 3) F. E. Low, Phys. Rev. **97** (1955), 1392.
- 4) L. Castillejo, R. H. Dalitz and F. J. Dyson, Phys. Rev. **101** (1956), 453.
- 5) G. F. Chew and F. E. Low, Phys. Rev. **101** (1956), 1570.
- 6) S. T. Ma, Rev. Mod. Phys. **25** (1953), 853.
H. Araki, Y. Munakata, M. Kawaguchi and T. Gotô, Prog. Theor. Phys. **17** (1957), 419.

Added note. After communication to the editors I became aware of the paper by Martin,* in which he had studied in detail the validity of Levinson's theorem for the scattering due to a non-local but separable potential and shown that this theorem is valid only when there is no degeneracy for positive energy wave functions.

Superconductivity of Ferromagnetic Metals

T. Kasuya

Physical Institute, Nagoya University

October 3, 1958

In rare earth metals, Lanthanum exhibits superconductivity with the transition tem-

* A. Martin, Nuovo Cimento, **7** (1958), 607.

perature $4 \sim 6^\circ \text{K}$. However, the other elements do not exhibit superconductivity. This phenomenon is explained by the recent theory of superconductivity by Bardeen et al.¹⁾ and the so-called $s-d$ interaction investigated by us.²⁾

The Hamiltonians of electron-phonon and $s-d$ interactions are written as

$$H_{ep} = \sum_k \sum_q \sum_\nu \{ C(q) a_{k+q,\nu}^* a_{k,\nu} b_q + C^*(q) a_{k,\nu}^* a_{k+q,\nu} b_q^* \} \quad (1)$$

$$H_{sd} = \sum_n \sum_k \sum_q (g-1) J(q) N^{-1} e^{iqR_n} \times \{ j_n^z (a_{k+q,+}^* a_{k,+} - a_{k+q,-}^* a_{k,-}) + j_n^+ a_{k+q,-}^* a_{k,+} + j_n^- a_{k,+}^* a_{k+q,-} \} \quad (2)$$

where $C(q)$ is the coupling constant of electron phonon interaction, $J(q)$ the exchange energy between conduction and unfilled inner shell electrons, $a_{k,\nu}^*$, $a_{k,\nu}$ the creation and annihilation operators of electron with wave vector k , spin ν and b_q^* , b_q the operators of longitudinal phonon with wave vector q .

For the nearly degenerated electronic states near the Fermi surface, the effective Hamiltonians of binary electrons interaction are deduced from the second order perturbation process as follows,

$$H_{ep}^{(2)} = - \sum_{k \neq k'} \sum_q \sum_\nu \sum_\mu V_{ep}(q) \times a_{k+q,\nu}^* a_{k,\nu} a_{k',\mu}^* a_{k'+q,\mu} \quad (3)$$

$$H_{sd}^{(2)} = - \sum_{k \neq k'} \sum_q V_{sd}(q) \times a_{k+q,-}^* a_{k,+} + a_{k',+}^* a_{k'+q,-} \quad (4)$$

In the calculation of $H_{sd}^{(2)}$, we used the molecular field approximation, and neglected the first term of eq. (2) because this term has the same character as that of ordinary impurity scattering, that is, the freedom of j_n is not concerned. Further, neglecting the energy differences between electronic

states in initial, final and intermediate states compared with the energies of phonons and the molecular field of j_n , we write V_{ep} and V_{sd} as

$$V_{ep}(q) = |C(q)|^2 / \hbar \omega_q \quad (5)$$

$$V_{sd}(q) = 2(g-1)^2 \bar{j}_n^z J^2(q) / N H_m \quad (6)$$

where ω_q is the angular frequency of the longitudinal phonon with wave vector q , \bar{j}_n^z the mean value of j_n^z , and H_m the molecular field.

As Bardeen et al. have shown, the important terms for the superconductivity are of the type of $a_{k,+}^* + a_{-k,-}^* - a_{-k',-} - a_{k',+}$. Then, considering the screened Coulomb interaction, the Bardeen type Hamiltonian becomes

$$H_0^{(2)} = - \sum_{k \neq k'} \sum_q \left(2V_{ep} - V_{sd} - \frac{4\pi e^2}{k_c^2 + |\mathbf{k} - \mathbf{k}'|^2} \right) a_{k,+}^* + a_{-k,-}^* - a_{-k',-} - a_{k',+} \quad (7)$$

As is seen from the above formula, $H_{sd}^{(2)}$ has the effect to diminish the effect of $H_{ep}^{(2)}$, and thus if the $s-d$ interaction is sufficiently large, the superconductivity is destroyed.

Now let us calculate the order of magnitude of each term in eq. (7). With the use of Bardeen's coupling constant of electron phonon interaction,³⁾ V_{ep} becomes

$$V_{ep} = \frac{0.22}{N} \frac{\zeta^2}{M u_0^2} \quad (8)$$

where ζ is the Fermi energy, M the mass of the ion and u_0 the velocity of the longitudinal phonon. For example, using the data of Gadolinium, that is,

$$N = 3.07 \times 10^{22} \text{ c.c}^{-1}, \quad \zeta = 1.2 \times 10^{-11} \text{ erg.}$$

$$M = 2.6 \times 10^{-22} \text{ gr.}$$

$$u_0 = 1.7 \times 10^5 \text{ cm/sec } (\theta = 160)$$

$$J(q) = J(0) = 2.5 \times 10^{-13} \text{ erg.}$$

$$H_m = \frac{3}{j(j+1)} \kappa T_p = 0.79 \times 10^{-14} \text{ erg.} \quad (9)$$

where we used the free electron approximation and assumed that n/N equals 3, we obtain

$$V_{ev} = 4.2 \times 10^{-12} \frac{1}{N} \text{ erg.}$$

$$V_{sd} = 5.5 \times 10^{-11} \frac{1}{N} \text{ erg.}$$

$$\frac{4\pi e^2}{k_c^2 + q^2} = \frac{4\pi e^2}{(2k_0)^2} = 1.1 \times 10^{-12} \frac{1}{N} \text{ erg.} \quad (10)$$

where k_0 is the wave vector at the Fermi surface.

Thus we may say that Gd does not exhibit superconductivity. Also in other elements of rare earth metals except La, the situations are nearly same as that of Gd and we may say that they do not exhibit superconductivity.

In transition elements of Fe-families such as Fe, Ni, etc., $s-d$ exchange integral $J(0)$ becomes much larger than that of rare earth metals and it is consistent with the fact that they do not exhibit superconductivity.

In transition elements much heavier than rare earth metals of La-families, $J(0)$ becomes smaller and thus in some elements there would appear superconductivity.

Finally, the author expresses his cordial thanks to Asso. Prof. S. Nakajima for valuable discussions.

- 1) J. Bardeen, L. N. Cooper and J. R. Schrieffer, *Phys. Rev.* **106** (1957), 162.
N. N. Bogolubov.
- 2) T. Kasuya, *Prog. Theor. Phys.* **16** (1956), 45.
- 3) J. Bardeen, *Phys. Rev.* **52** (1937), 688.

Proton-Proton Scattering in the Gev Region

—Criticism of the Blockhintsev-Bubelev Model—

Junko Ōba* and Tetsuo Gotō**

*Department of Physics, St. Paul's University

**Department of Physics, College of Science and Engineering, Nihon University

July 28, 1958

According to the recent development of experimental analyses in the Gev region, "the nucleon structure" has become an important problem and the existence of "the nucleon core" has been indicated by the experiments.

Blockhintsev and Bubelev dealt with nucleon-nucleon collisions using the Williams-Weizsäcker method and in their treatment the central collision was considered as a matter of little importance.

Now we want to examine the Blockhintsev-Bubelev model in detail. That is, we compute, not only the total inelastic cross section, but also the cross section producing two or more pions. In this case, the inelastic cross section rapidly changes with the lower limit of the impact parameter, i. e. the cut-off parameter and with the coupling constant. When Bubelev's cut-off parameter is used the branching ratio of the production cross sections of one pion to two or more pions turns out to be larger than the experimental value. Moreover, we are able to calculate the elastic cross section from the inelastic cross section by making use of the impact parameter in this model. Here we used the assumption that Itō and Minami postulated, i. e. the most part of the elastic scattering was due

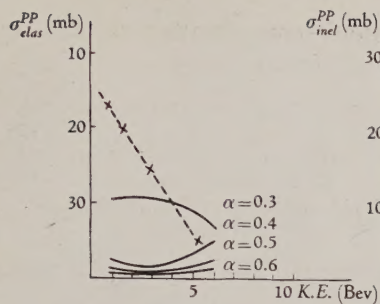


Fig. 1. The elastic cross section.

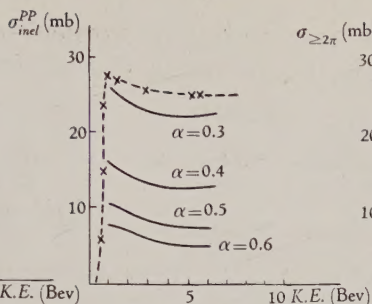


Fig. 2. The inelastic cross section.

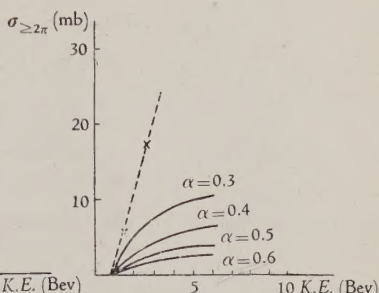


Fig. 3. The cross section for two or more pions production.

to the shadow scattering. Using the former cut-off parameter, we obtain the result that the elastic cross section also gives a much smaller value than the experimental value. Thus, the fact that the experiment gives a larger value than the computed value shows that this cut-off parameter is not suitable as long as we take $g^2/\hbar c \sim 15$. Though the validity of the Williams-Weizsäcker method or the calculation method of clouds of pions is open to question, we do not discuss these problems now. Thus σ_{elas} , σ_{inel} , and $\sigma_{\geq 2\pi}$ which correspond to various cut-off parameters are estimated respectively. The result is shown in Figs. 1~3.

The experimental results are shown by the dotted line. In the figure, $\alpha = r_c/R$, where R is $\hbar/m_\pi c$ and r_c is the lower limit of the impact parameter. Judging from the above figures, the inelastic total cross section appears to be explained with Bubelev's cut-off parameter, where $\alpha = 0.3$ corresponds to Bubelev's cut-off parameter. However, we

cannot understand at all the elastic cross section and the production cross section of two or more pions with $\alpha = 0.3$. Therefore, contrary to Bubelev's model, we do not consider that the contribution from the central collision is negligible, and it is taken as an adjustable parameter. By taking into account the contribution from the central collision, of which the maximum value is the geometrical cross section, the most compatible cut-off parameter amounts to about twice Bubelev's cut-off parameter.

Thus, when we investigate many processes in detail, it is doubtful that nucleon-nucleon collisions are able to be comprehended only by peripheral collisions. It seems to be reasonable that the central collisions participate in nucleon-nucleon collisions in the Gev region.

- 1) D. Blockhintsev, the CERN Symposium, JETP, 32 (1957), 350.
- 2) E. G. Bubelev, JETP, 33 (1957), 539.
- 3) D. Ito and S. Minami, Soryushiron-Kenkyu, (1957), 1, 27.

Errata

Dispersion Relations in Nucleon-Nucleon Scattering Prog. Theor. Phys. 19 (1958), 517.
Sadahiko MATSUYAMA and Hironari MIYAZAWA

Equations (6.4) and (7.1) should read

$$F(\omega) \sim -\frac{8\pi}{m^2} \frac{R^{-1}}{1-R^{-1}r} \frac{1}{\omega-\omega_D}$$

and

$$D_0(s)(\omega) = -U(s)(\omega) + \frac{p}{\pi} \int_m^\infty \frac{A_0(s)(\omega')}{\omega' - \omega} d\omega' - \delta_{SD} \left(\frac{8\pi}{m^2} \frac{R^{-1}}{1-R^{-1}r} \frac{1}{\omega-\omega_D} \right) + \text{const} \quad (7.1)$$

respectively.

JOURNAL OF THE PHYSICAL SOCIETY OF JAPAN

Volume 13, Number 10, October 1958

CONTENTS

	Page
Excited States of Pt^{192} Mitsuhiro KAWAMURA, Atsushi AOKI and Takeo HAYASHI	1071
Sector Type Double Focusing Beta Ray Spectrometer. Part I. General Properties Mitsuo SAKAI and Hidetsugu Ikegami	1076
Sector Type Double Focusing Beta-Ray Spectrometer. Part II. Twin Setting Measure- ment of the Weak Higher Component of Cs^{137} Toshimitsu YAMAZAKI, Hidetsugu Ikegami and Mitsuo SAKAI	1080
Temperature Distribution during the Martensite Plate Formation Zenji NISHIYAMA, Akira Tsubaki, Hideo Suzuki and Yasusada Yamada	1084
Theory of Thermoelectric Power of Ionic Crystals (I) Eijirō HAGA	1090
A Theory of Impurity Conduction. I Tadao KASUYA	1096
Magnetic Property and Crystal Chemistry of Ilmenite ($MeTiO_3$) and Hematite (αFe_2O_3) System. I. Crystal Chemistry Yoshikazu ISHIKAWA and Syun-iti AKIMOTO	1110
Lattice Defects in Zincblende. Part I. Phenomenological Expressions of Interionic Potentials Sumitada ASANO and Yasuo TOMISHIMA	1119
Lattice Defects in Zincblende. Part II. Formation Energy of Lattice Defects Sumitada ASANO and Yasuo TOMISHIMA	1126
Internal Friction of Lead. I. Yosio HIKI	1138
Scattering of Electrons by Carbon Films Yoshihiro KAMIYA	1144
The Crystal Structure of Germanium Selenide $GeSe$ Atsushi OKAZAKI	1151
Anisotropy of Nuclear Magnetic Resonance due to Crystal Dislocations. Eizo OTSUKA	1155
Effect of Velocity Gradient on the Radial Distribution of Polymer Molecules in Dilute Solution Ryuzo KOYAMA	1167
Paramagnetic Resonance Spectrum of Mn^{++} in NaCl Single Crystal Kazuo MORIGAKI, Minoru FUJIMOTO and Junkichi ITOH	1174
Nuclear Magnetic Resonance Studies of Iodine Ions in Aqueous Solution Junkichi ITOH and Yukio YAMAGATA	1182
The Lower Excited States and the Phosphorescent State of Diphenyl. II. Scheme of Modification of MO Kaoru IGUCHI	1186
Methods of Approximation of Molecular Integrals. I Sigeru HUZINAGA	1189
The Visco-elastic Properties of Network Structure. III. Normal Stress Effect (Weis- senberg Effect) Misazo YAMAMOTO	1200
Disturbance Phenomena in Probe Measurement of Ionized Gases Takayoshi OKUDA and Kenzo YAMAMOTO	1212
Magneto-hydrodynamic Oscillations of a Perfectly Conducting Fluid Sphere Placed in a Uniform Magnetic Field Tsuneji RIKITAKE	1224

SHORT NOTES

Chemical and Thermal Polish of KCl Crystals Masayoshi SAKAMOTO and Shunsuke KOBAYASHI	1231
Nuclear Magnetic Resonance Experiments of I^{127} in KI Junkichi ITOH and Yukio YAMAGATA	1232
A Proposed Explanation of Irradiation Growth of Uranium Ryukiti R. HASHIGUTI, Hideo SAKAIRI and Tetuya SUGAI	1233
On the Conduction Mechanism of Oxide Coated Cathode. Hiroyuki MIZUNO	1234
The Far Infrared Absorption of Sulfides, Selenides and Tellurides of Zinc and Cadmium Akiyoshi MITSUISHI, Hiroshi YOSHINAGA and Shigeru FUJITA	1235
The Far Infrared Absorption of Ferrites Akiyoshi MITSUISHI, Hiroshi YOSHINAGA and Shigeru FUJITA	1236
Growth and Extinction of Guanidine Aluminum Sulfate Hexahydrate Crystals Atsuo NISHIOKA and Kyoza SEKIKAWA	1237

Errata

The Electron Distribution and the Location of Bonded Hydrogen Atom in Crystals. Yujiro TOMIIE	1239
Diffusion of Cobalt into Single Crystals of Copper Masanobu SAKAMOTO	1239
Piezo-Effect of Selenium Rectifiers Tadasu NUMATA	1239